

KM	tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
KJ	chemotaxis; chemokinesis; hemostrasis; gene therapy; thrombolysis;
KI	receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
XX	
OS	Homo sapiens.
XN	MO9845437-A2.
PN	
PD	15-OCT-1998.
PF	10-APR-1998; 98MO-US06956.
PR	10-APR-1997; 97US-0837312.
PA	(GEMV ) GENETICS INST. INC.
PI	Agostino MJ, Jacobs K, Lavallie ER, McFoy JM, Merberg D;
PT	Macle LA, Spaulding V, Treacy M;
PS	WPI 1999-07q078/06.
XX	
DR	New polynucleotides encoding human secreted proteins - derived from e.g. human blood, kidney, fetal lung, placenta, testes, brain, ovary, pituitary, retina and colon cDNA libraries.
XX	
PS	Claim 1: Page 178; 641pp; English.
CC	The present sequence represents an expressed sequence tag (EST), and is a polynucleotide of the invention. The polynucleotides of the invention are all secreted EST sequences and proteins isolated from a variety of human tissue sources. The EST sequences and proteins encoded by them are predicted to have useful biological activities which would make them suitable for treating, preventing or ameliorating medical conditions in humans and animals, although no supporting data is given. Suggested activities include nutritional activity, immune stimulating or suppressing activity, haemopoiesis regulating activity, tissue growth activity, activin/inhibin activity, chemotactic/chemokinetic activity, hemostatic and thrombolytic activity, receptor/ligand activity, anti-inflammatory activity, cadherin/tumour invasion suppressor activity, tumour inhibition activity. The EST sequences are also stated to be useful for gene therapy.
CC	
XX	
SQ	Sequence 686 BP; 137 A; 243 C; 177 G; 129 T; 0 other;
Alignment Scores:	
Pred. No. :	0 Length: 686
Score: 1107.80 Matches: 218	
Percent Similarity: 98.64% Conservative: 0	
Best Local Similarity: 98.64% Mismatches: 2	
Query Match: 61.44% Indels: 2	
DB: 2 Gaps: 0	
us-09-831-455b-6 (1-347) x AAV87784 (1-686)	
OY 1 MetSerSerTrrPserArgInArProLySerProGlyIleGlnProHisValSer 20	
Db 27 ATGTCTTCCTTGAGCAGACGACGACCAAAAAGGCCAGGGGGCATTCACCCCAGTTCTT 86	
OY 21 ArgThrLeuPheLeuLeuLeuLeuLeuAlaISerLAsTrGcLYalThLeuSerPro 40	
Db 87 AGAACTCGTCTCGCTGCTGCTGTGTGGCAACCCKAACCCTGGGGGGGTACCTTAGCCCC 146	
OY 41 LysApSCySgInValPheArgSerAspHisGLYSerSerILeserCySgInProProAla 60	
Db 147 AAAGACTGCAGCGATGTTCCGCTCAGAACACGACGACGACGTCATCTCTGTCAACACCGCC 206	
OY 61 GluileProGlyITrLeuProAlaSPTrValHISLeuAlaValGlnPHePheAsnLeu 80	
Db 207 GAATCCCGCGGCTACTCGCACGACGACACCGCTGACCTGGCGGTGATTCTTAACCTG 266	
OY 81 ThrHisLeuProAlaAsnLeuLeuGlnGlyALIsertLyLeuGlnIleuHisLeuSer 100	
Db 267 ACCCACTGCACACCAACTCTCTCAGGGGGGCGCTTAACCTCCAAGAATTGCACTTCTCC 326	

QY 101 SeranGlyLeuGluSerLeuSerProGluPheLeuArgProValProGluLeuArgVal 120  
 DB 327 AGCAATGGGGCTGGAAACCCCTCTCGCCGAATTCCTGCGGCAAGTGGCCAGCTGAGGGTG 386  
 QY 121 LeuAspLeuThrArgAsnAlaLeuThrGlyLeuProProGlyLeuPheGlnAlaSerAla 140  
 DB 387 CTGGATCTAAACCCGAAACGCCCTGACCGGGCTGGCCCTCTCTCCAGGCTCAGCC 446  
 QY 447 LeuAspThrLeuValLeuGluSerGluAsnGlnLeuGluValLeuGluValSerTyrLeu 160  
 DB 447 ACCCTGGACACCCCTGGATTAAGAAACCCAGCTGAGGCTCTGGAGGCTCTGTGGCTA 506  
 QY 160 HisGlyLeuLysAlaLeuGlyHisLeuAspLeuSerGlyAsnArgLeuArgLysLeuPro 180  
 DB 537 CAGGCGCTGAAAGCTCTGGGCGATCTGGACCTGTGGAAACGCCCTCGGAAACCTGCC 566  
 QY 537 ThrGlyLeuLeuAlaAspPheThrLeuThrLeuAspLeuGlyLysHisGlnLeu 200  
 DB 567 CCGGGCTGCTGGCCAACTTCACTCTCTGCGACCCCTGACCTGGGGAGAACCGATTG 626  
 QY 201 GluThrLeuProProAspLeuLeuArgGlyProLeuGlnLeuGluArgLysLeu 220  
 DB 627 GAGACCTTGGACCTGACCTCTGAGGGGTCCGTCGCAATTAGACGGCA-CATCTCTGA 685  
 QY 220 u 220  
 DB 686 G 686  
 DE EST clone ES35.  
 XX Expressed sequence tag; secreted protein; haematopoiesis regulator;  
 KW tissue growth; activin; inhibitor; tumour invasion suppressor; EST; human;  
 KW chemokinesis; chemokinesis; haemostasis; gene therapy; thrombolysis;  
 KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.  
 OS Homo sapiens.  
 PN W09845437-A2.  
 PD 15-OCT-1998.  
 PF 10-APR-1998; 98MO-US06956.  
 PR 10-APR-1997; 97US-0837312.  
 PA (GENY) GENETICS INST INC.  
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;  
 PI Racie LA, Spaulding V, Treacy M.  
 DR MPI: 1999-070078/06.  
 PT New polynucleotides encoding human secreted proteins - derived from  
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,  
 PT ovary, pituitary, retina and colon cDNA libraries  
 XX  
 PS Claim 1; Page 178; 641pp; English.  
 CC The present sequence represents an expressed sequence tag (EST), and is  
 CC a polynucleotide of the invention. The polynucleotides of the invention  
 CC are all secreted EST sequences isolated from a variety of human tissue  
 CC sources. The EST sequences and proteins encoded by them are predicted to  
 CC have useful biological activities which would make them suitable for  
 CC treating, preventing or ameliorating medical conditions in humans and

CC animals, although no supporting data is given. Suggested activities  
 CC include nutritional activity, immune stimulating or suppressing activity,  
 CC haematopoiesis regulating activity, tissue growth activity,  
 CC activin/inhibitor activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition  
 CC therapy. The EST sequences are also stated to be useful for gene  
 XX  
 SX Sequence 686 BP; 137 A; 243 C; 177 G; 129 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0 Length: 686  
 Score: 165.40 Matches: 88  
 Percent Similarity: 23.718 Conservative: 22  
 Best Local Similarity: 18.978 Mismatches: 62  
 Query Match: 9.17% Indels: 301  
 DB: 2 Gaps: 7  
 us-09-8455b-6 (1-347) x AAV87784 (1-686)  
 QY 27 LeuLeuDeuAlaSerAlaTrrpGlyValThrLeuSerProLysAspGlnValPhe 46  
 DB 672 GTTCTAATTCACAGC-----GACCCCTCAGAGAGTCAAGTGGCA 634  
 QY 47 ArgSerAspHisGlySer-----SerLysSerCys----- 56  
 DB 633 AGGCTCCCACTGGCTTCTCCCAAGGTCA--GGGTCCGAGAGGAGTGAATGGCCAG 576  
 QY 57 GluProProAlaGluThrProGlyTyrLeuProAlaAspThrValHisLeuAlaValGlu 76  
 DB 575 CAGCCCGGGGGGAGATTCCGAGCGGTTCCA--GACAG-GTCCAGAT-GCCCCAGAG 520  
 QY 77 PhePheAsnLeuThrHisLeuProAlaAsnLeuGlnGly-----AlaSerLys-Le 94  
 DB 519 -----CTTTCAGCGCCGTAGCACAGAGACT 493  
 QY 94 uGlnGluLeuHisLeu-----SerSerAsn-----GlyLeuGlu 105  
 DB 492 CCAAGACCTCCAGCTGGTTCTTCAATCCAGGGGTGCCAGGGGTGAGCCCTGGA 433  
 QY 105 uSerLeu-----SerPro-----GluPhe-----LeuArgProVal-ProGlnLeuArg 119  
 DB 432 AGAGGCCCGAGGCGAGCCGGCTCAGGGCTGTAGATTCAGCACCTCAGCTGCGG 373  
 QY 119 gValLeuAspLeuThrArgAsnAlaLeuThrGlyLeuProProGlyLeuPheGlnAlaLe 139  
 DB 372 G----- 372  
 QY 139 fAlaThrLeuAspThrLeuValLeuLysGluAsnGlnLeuGluValLeuGluValSerT 159  
 DB 371 -----CACTG 367  
 QY 159 pLeuHisGlyLeuLysAlaLeuGlyHisLeuAspLeuSerGlyAsnArgLeuArgLysLe 179  
 DB 366 GCC-GCAGGAATTCGGGG----- 349  
 QY 179 uProProGlyLeuLeuAlaAspPheThrLeuLeuArgThrLeuAspLeuGlyGluAsnG 199  
 DB 349 ----- 349  
 QY 199 nLeuGluThrLeuProProAspLeuLeuArgGly-----ProLeuGlnLeuGluArg 216  
 DB 348 -----AAAGGCTTCCAGGCCCATTCG-CTGGAGAGAGT 319  
 QY 217 -----LeuHisLeuGlnGly-----AsnLysLeu-----GlnValLeuGlyLysAs 230  
 DB 318 GCMAATCTTGAGCTTAGAGCGCCCTGAGGAGAGGTGGCTGGACGGTGG----- 270  
 QY 230 pLeuLeuLeuProGlnProAspLeuArgTyrLeuPheLeuAsnGly-----AsnL 247  
 DB 269 -----GCTCAGGTGAAGAAATT 253

QY 247 ysluAlaArg-----ValAlaAlaGly-AlaPheGlnGlyLeuArgGln---LeuAsp 263  
 DB 232 CCAGGCGAGTGCAGCGTGTGCTGCGTGCAGGTGCGGCGGATTCGGCGAGGCTGAC 193  
 QY 264 MetleuAspLeuSerAsnSerLeuAlaSerValProGlnGlyLeuTrpAlaSerLeu 283  
 DB 192 AGCA-IGATGCG-AGTGCATGCTGTGACCGAACCCTGCTTGGGCTGAC-- 138  
 QY 284 GtlyGlnPro-----AsnTrpAspMetArgAsp-GlyPhe---AspIle 296  
 DB 137 GGTACACCCCGCAGCTGAGCTGCCACGACGACGAGACAGAGATTCGTGAACATG 78  
 QY 297 SerGlyAsn---ProTrpIleCysAspGlnAsnLeuSerAspLeuTrpArgTrp---Leu 314  
 DB 77 GGGTTGAATGCCCCCGGCTTTT-----MetPheSerGln 324  
 QY 315 GtlyGlnGtlyGlnAspGly-----MetPheSerGln 324  
 DB 44 GGTCTCCAGAGACATGATGATGAGGCTCTTGGCCGA 5  
 RESULT 3  
 AA622495 403 bp mRNA linear EST 14-OCT-1997  
 LOCUS n959B04.s1 NCI-CGAP-CO9 Homo sapiens CDNA clone IMAGE:1148143 3'  
 DEFINITION similar to contains Alu repetitive element; mRNA sequence.  
 AA622495  
 VERSION AA622495.1 GI:2526371  
 KEYWORDS EST  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-remail.nih.gov  
 Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D.,  
 Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www.bio.llnl.gov/dbp/image/image.html  
 Seq primer: -40m13 fwd. ET from Amerisham.  
 Location/Qualifiers  
 1. 403  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1148143"  
 /tissue\_type="colon tumor RER+"  
 /lab\_host="DH10B"  
 /clone\_id="NCI-CGAP-CO9"  
 /note="Organ: colon; Vector: pUTR3D-Pac (Pharmacia) with a  
 modified polylinker; 1st strand cDNA was prepared from  
 RER+ colon tumor, and was then primed with a Not I -  
 oligo(dt) primer. Double-stranded cDNA was ligated to Eco  
 RI adaptors (Pharmacia), digested with Not I and cloned  
 into the Not I and Eco RI sites of the modified pUTR3  
 vector. Library is not normalized. Library was  
 constructed by Bento Soares and M. Fatima Bonaldo (Soares4)  
 )"  
 BASE COUNT 105 a 89 c 117 g 92 t  
 Alignment Scores:  
 Pred. No.: 0 Length: 403  
 Score: 116.40 Matches: 45  
 Percent Similarity: 20.69% Conservative: 21  
 Best Local Similarity: 14.11% Mismatches: 29

Query Match: 6.46% Indels: 225  
 DB: 1 Gaps: 6  
 us -09-831-455b-6 (1-347) x AA622495 (1-403)  
 QY 134 GtlyLeuPheGlnAlaSerAlaThrLeuAspPheLeuValLeuGlnGlnGlnGln 153  
 DB 13 GGTATTATT----- 21  
 QY 154 ValLeuGlnValSerTrpLeuHisGlyLeuValAlaLeuGlnHisLeu---AspLeu 172  
 DB 22 -----ATTAAACCCCTTGCTT---CTGACCCCAAGCTAA 52  
 QY 172 ergtlyAsnArgLeuArgGlyLeu-----ProProGlyLeuLeuAlaAsnPheth 188  
 DB 53 GTGGAGCTCAGCATCGACCTGACCTGACAGACCCCT----- 92  
 QY 188 rleuLeuArgTrpThrLeuAspLeuGlyGlnAsnGlnLeuGlnThrLeuProProAspLeu 208  
 DB 93 -----GACTTGGGAGACAGAGTCAAG----- 116  
 QY 208 uargtlyProLeuGlnLeuGlnArgLeuHisLeuGlnGly---AsnIleLeuGlnValLeu 227  
 DB 117 -AAAGCAGTCCCGACAGGCGGCGGCGGCTGACGAGGAGAGACAG----- 163  
 QY 228 GtlyAspLeuLeuLeuProGlnProAspLeuArgTyrLeuPheLeuAsnGlyAsnIle 247  
 DB 163 ----- 163  
 QY 248 LeuAlaArgValAlaAlaGlyAlaPheGlnGlyLeuArgGlnLeuAspMetLeuAspLeu 267  
 DB 163 ----- 163  
 QY 268 SerAsnAsnSerLeuAlaSerValProGlnGlyLeuTrpAlaSerLeuGlnGlnProAsn 287  
 DB 164 -----GGCTGGAGAAACCCAGA 181  
 QY 288 -----TrpAspMetArgAspGly-PheAspIleSerGlyAsnProTrp----- 301  
 DB 182 GTTCAGAGATCTGGAAACAGAGACGATTCAGACGAGGACACACCTTGGCATG 241  
 QY 302 -----IleCysAspGlnAsnLeuSerAspLeuTrpArgTrpLeuGln- 315  
 DB 242 GGTAGAGAAAGCCCATCTGCTGTCTAGAAC-----CATGATGATGAAC 289  
 QY 316 -----AlaGlnIleAspGlyMetPheSerGlnAsnAspThrArgCysAlaGlyPro 332  
 DB 290 GGGGTCCAGCCGCGTATGAC-----AGCCCC 316  
 QY 333 Gtly-----AlaValIleGly-----Gln 338  
 DB 317 GACTTCACCTGTAAAGACGATATTGCTTTTAATTAAATTAAATTTTGGAG 376  
 QY 339 ThrLeuLeu-AlaValAlaIleSer 346  
 DB 377 ACAGTCTTCTCTGTGCGCCAGGCT 401  
 RESULT 4  
 AA622495/c 403 bp mRNA linear EST 14-OCT-1997  
 LOCUS n959B04.s1 NCI-CGAP-CO9 Homo sapiens CDNA clone IMAGE:1148143 3'  
 DEFINITION similar to contains Alu repetitive element; mRNA sequence.  
 AA622495  
 VERSION AA622495.1 GI:2526371  
 KEYWORDS EST  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index

## JOURNAL COMMENT

Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: e9apbs@mail.nih.gov  
Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone Distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
www.bio.llnl.gov/bbrp/image/image.html  
Seq primer: -40m13 fwd. RT from Amersham.  
Location/Qualifiers

## FEATURES

1. 403  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_image="1148143"  
/tissue\_type="colon tumor RER+"  
/lab\_host="DH10B"  
/clone\_id="NCI-CCAP\_Co9"  
/note="Organ: colon; Vector: p773D-pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from RER+ colon tumor, and was then primed with a Not I - Oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library is not normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo (Soares4)

BASE COUNT 105 a 89 c 117 g 92 t

## Alignment Scores:

Pred. No.: 0 Length: 403  
Score: 98.00 Matches: 53  
Percent Similarity: 24.52% Conservative: 11  
Best Local Similarity: 20.31% Mismatches: 41  
Query Match: 5.44% Indels: 158  
DB: 1 Gaps: 4

us-09-831-455b-6 (1-347) x AA622495 (1-403)

```

OY 84 ProAlaAsnLeuLeuGlnGlyAlaSerLysLeuGlnGluLeuHisLeuSerSerAsn--- 102
Db 399 CCGGGCGGACAGACAGACAGCTGCTCAAAATTAATAATTAATAAACAATTAAT 340
OY 103 -----GlyLeuGlu-----SerLeuSerProGluPheLeuAr 113
Db 339 CCGCCTTTACAGGTGAAGTGGGCTGTCATAGCGGCTGGAGCCCGTTTCATCA 280
OY 113 gPro--ValProGln--LeuArgValLeuAspLeuThr-----ArgAsnAlaLeu-- 128
Db 279 TCCATGCTTCTAGAACACACAGATGGGCTTCTTACCATGCCCAAGGTGCTCCG 220
OY 129 ThrGlyLeuProGlyLeuPhe--GlnAlaSerAlaThrLeuAspThrLeuValLeuL 148
Db 219 TCTGGAATGCCGT--TCCGTGTTCCAGATCTCTGAACCTCGGGTCT----- 172
OY 148 ySGluAsnGlnLeuGluValLeuGlnValSerTrpLeuHisGlyLeuYsaAlaLeuGlyH 168
Db 172 ----- 172
OY 168 IsLeuAspLeuSerGlyAsnArgLeuArgLysLeuProProGlyLeuLeuAla--AsnPhe 187
Db 171 -----CCTACCCCTTG--CTTCTCTG 150
OY 186 TrpLeu-----LeuArgThrLeuAspLeuGlyGluAsnGlnLeuGluThrLeuProPro 205
Db 149 CAGCTGAGCCCTGCGCACACTGGGCTG-----CCTTT 117
OY 206 ---AspLeuLeuArgLysProLeuGlnLeuGlnArgLeuHisLeuGlnGlyAsnLysLe 224
Db 116 CTGGGACT--CTGCTTCTCCCAAGTCAG----- 91

```

```

OY 224 uGlnValLeuGlnGlyAspLeuLeuLeuProGlnProAspLeuArgTrpLeuPheLeuAs 244
Db 91 ----- 91
OY 244 nGlyAsnLysLeuAlaArgValAlaAlaGlyAlaPhe-----GlnGlyLeuArgGlnLe 262
Db 90 -----GGGGCTCTCTGAGTGCAGGCGTCTGARG----- 64
OY 262 uAspMetLeuAspLeuSerAsnAsn--SerLeuAlaSerValProGluGlyLeu 279
Db 63 -----CTGAGTCCCACTTACCTTGGGGTCTGAGACCAAGGCTTTA 24

```

Search completed: September 9, 2003, 15:54:01  
Job time : 1 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2003, 15:44:23 ; Search time 0.001 seconds  
(without alignments)  
114.857 Million cell updates/sec

Title: us-09-831-455b-6  
Perfect score: 1803  
Sequence: 1 MSSWSRQRPRKSPGIGPHVS.....RCAGPEAVKQOTLLAVAKSQ 347

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 331 residues

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : baa31994.genpept.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	434.5	24.1	331 1 BAA31994	

ALIGNMENTS

RESULT 1  
BAA31994

Query Match 24.1%; Score 434.5; DB 1; Length 331;  
Best Local Similarity 36.0%; Pred. No. 0;  
Matches 107; Conservative 46; Mismatches 129; Indels 15; Gaps 4;

QY	56	COPRA--EIPGYLADYVHLAVEFPNLTLPANLQAGSKLOELHLSNGLESPEFLR 113
DB	41	CNSPSLHEFPPTGFPARAKMISVEFTQVSSLGVEALQGLPNQELHLSNNRKLTPSGGLR 100
QY	114	PVPOLRVLDLTRNALTGPGIFQASATLDTLVLEKNOLEVEVSWLHGLKALGHLDLSG 173
DB	101	NLPOLHTLDLSTNLELDPPEITFNASSLILPLSENOLAEIHPSWFOTLGEIRLGLDH 160
QY	174	NRLRKLPPGLIANFTLLFTLIDLGENDLETPDLLRGPLQLEERLLEGNKLOVLGKDLL 233
DB	161	NOVKEIPISCFDKLKLTLSDLSFNLRRLAPEMPSGIDNLEKILLESNPICIVGRFFH 220
QY	234	PQPLRLTYFLNGNKLARVAGAFQGLRQDMLDLSNNSLASVPEGIM---ASLGOPNMDM 290
DB	221	WPKLTYLISLNSSITFNI-MGFQPLRQLDLDSDNELTTMEPPVYKTSANL----- 272
QY	291	RDGFDISGNPWICQNTSLDYRMLOAQDKMFSONDFRCAGPEAVKQOTLLAVAKSQ 347
DB	273	--SLDLSGNPWACDCRLDNLTWYNEINIHLYSKKEIYCASPKHFKGCATSLHKSQ 327

Search completed: September 9, 2003, 15:44:23  
Job time : 0.001 secs

!!AA\_SEQUENCE 1.0  
BAA31994 GB:AB007198 phospholipase A2 inhibitor [Agkistrodon blomhoffii  
sinuatus] (ver 1)

BAA31994 Length: 331 September 9, 2003 15:35 Type: P Check: 2681 ..

```
1  MKSSVPSLLI ACLVMSLSY TQGVLCPT PAPENVTEFV CNSPSLHEFP
51  TGFPAKMI SVEFTQYSSL GVEALQGLPN LQELHLSNMR IKTLPSSGLFR
101 NLPQHTIDL STNHEDLPP EFTNASSLI LPLSENOLA ELHPSWFQTL
151 GELRIIGLDH NQVKEIPISC FDKIKKLTSL DLSFNLRRL APEMFGLDN
201 LEKLIESNP IQCIVGRTFH WHPKLVYLSL KNSSLTNMG FQQLPQLLEL
251 IDLSNDELTT MEPPVYKTS A NLSLDLSGNP WACDCRLDNL LTVVNEHNH
301 LYSKEEIVCA SPKHFGGCA TSLHKSQICP C
```

## Inhibition of phospholipase A<sub>2</sub>

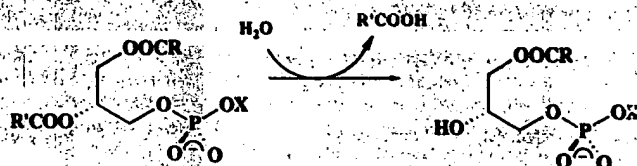
MICHAEL H. GELB,<sup>\*,1</sup> MAHENDRA K. JAIN,<sup>1</sup> AND OTTO G. BERG<sup>2</sup>

<sup>\*</sup>Departments of Chemistry and Biochemistry, University of Washington, Seattle, Washington 98195, USA; <sup>1</sup>Department of Chemistry and Biochemistry, University of Delaware, Newark, Delaware 19716, USA; and <sup>2</sup>Department of Molecular Biology, Uppsala University Biomedical Center, Uppsala, Sweden

**ABSTRACT** Phospholipases A<sub>2</sub> are involved in inflammatory processes such as the liberation of free arachidonic acid from the membrane pool for the biosynthesis of eicosanoids. Inhibitors of these enzymes are proving useful in determining the biological roles of phospholipases A<sub>2</sub> in complex cellular processes and may also have therapeutic potential. Inhibition of these lipolytic enzymes is more difficult to characterize as the enzymatic reaction occurs at a lipid/water interface. This review focuses on the description of a number of classes of rationally designed phospholipase A<sub>2</sub> inhibitors. The development of a theoretical framework for the proper analysis of inhibitors is presented. Structural studies of phospholipase A<sub>2</sub>-inhibitor complexes suggest how the lipolysis reaction is catalyzed. Finally, some recent results on the use of phospholipase A<sub>2</sub> inhibitors in living cells and tissues are revealed.—Gelb, M. H., Jain, M. K., Berg, O. G. Inhibition of phospholipase A<sub>2</sub>. *FASEB J.* 8: 916-924; 1994.

**Key Words:** lipases • inhibitors • arachidonic acid • eicosanoids • protein-lipid interactions

PHOSPHOLIPASES A<sub>2</sub> (PLA<sub>2</sub>s)<sup>2</sup> CATALYZE THE hydrolysis of the ester linkage at the *sn*-2 position of phospholipids to produce free fatty acids and lysophospholipids (Scheme 1) (1).



Scheme 1

The secreted enzymes (sPLA<sub>2</sub>s) from pancreas (2), venoms (3), and inflammatory cells and exudates (4, 5) are the best characterized. These enzymes are small (molecular mass near 14 kDa), require millimolar amounts of Ca<sup>2+</sup> as a catalytic cofactor but not for binding of the enzyme to membranes (6), and are related both in terms of primary (7) and three-dimensional structure (8-15). The sPLA<sub>2</sub>s display almost no preference for different types of naturally occurring phospholipid head groups (substituent X in Scheme 1) or the structure of the fatty acyl chains attached to the glycerol backbone (16).

Recently, a high molecular mass (87 kDa) cytosolic phospholipase A<sub>2</sub> (cPLA<sub>2</sub>) has been found in human, bovine, and rabbit platelets (17-19), rodent macrophage cell lines (20, 21, 22), human monocytic cell lines (U937 and THP1) (23-26), and rat kidney (27). In contrast to the sPLA<sub>2</sub>s, these cPLA<sub>2</sub>s are activated by submicromolar amounts of Ca<sup>2+</sup> and

preferentially hydrolyze phospholipids that have arachidonic acid at the *sn*-2 position. The cPLA<sub>2</sub> is found in the cytosol of cells in the absence of Ca<sup>2+</sup>, and it translocates to the particulate fraction in the presence of submicromolar amounts of Ca<sup>2+</sup> (23, 24, 26-30). Calcium also promotes the binding of the enzyme to phospholipid vesicles in vitro (31, 32). Isolation of the cDNA encoding the cPLA<sub>2</sub> from U937 cells indicates that the enzyme has no regions that are homologous to the sequences of sPLA<sub>2</sub>s (23, 33). Furthermore, the amino acid sequence reveals a stretch of 45 residues in the amino-terminal region that shows homology to Ca<sup>2+</sup>-dependent forms of protein kinase C and other Ca<sup>2+</sup>-dependent membrane binding proteins. A 140 amino acid fragment of this enzyme that contains the amino terminus was shown to bind to membranes in the presence of submicromolar amounts of Ca<sup>2+</sup> (23).

Gross (34) and co-workers have identified a novel Ca<sup>2+</sup>-independent PLA<sub>2</sub> found in myocardial cytosol. This enzyme prefers plasmalogens that contain arachidonic acid at the *sn*-2 position. Bovine brain cytosol also contains Ca<sup>2+</sup>-independent PLA<sub>2</sub> (35).

Undoubtedly other PLA<sub>2</sub>s in cells are involved in phospholipid remodeling and catabolism and in the degradation of phospholipids with oxidatively damaged fatty acyl chains (36).

The sPLA<sub>2</sub> found in inflammatory cells and exudates is almost certainly involved in inflammatory processes (see, for example, refs 37-42). This enzyme is often called the non-pancreatic secreted sPLA<sub>2</sub> because its amino acid sequence is more closely related to those of the sPLA<sub>2</sub>s in the venoms from crotalids and viperids (type II enzymes) rather than to those from pancreas (type I enzymes) (7). This enzyme is secreted from activated platelets, mast cells, and other cell types in response to agonists. The role of this enzyme in the degradation of foreign bacteria appears certain (39), its role in mast cell histamine release and T cell activation is being clarified (43, 44), and its involvement in liberating arachidonic acid from the membrane in, for example, a macrophage-derived cell line and endothelial cells has recently been suggested (45, 46). Because arachidonate metabolites are potent mediators of allergic and inflammatory reactions, PLA<sub>2</sub> inhibitors are expected to have therapeutic value. The pancreatic sPLA<sub>2</sub>, also called a type I sPLA<sub>2</sub>, is found in tissues other than those of digestive function and appears to have a proinflammatory role (47). This enzyme is thought to bind to cells via a specific receptor lead-

<sup>1</sup>To whom correspondence should be addressed, at: Department of Chemistry, BG-10, University of Washington, Seattle, WA 98195, USA.

<sup>2</sup>Abbreviations: PLA<sub>2</sub>, phospholipase A<sub>2</sub>; cPLA<sub>2</sub>, 87-kDa calcium-dependent cytosolic PLA<sub>2</sub> found in mammalian cells; sPLA<sub>2</sub>, 14-kDa secreted calcium-dependent PLA<sub>2</sub> found in venoms, digestive fluids, and mammalian inflammatory exudates and cells.

ing to the induction of cyclooxygenase and the subsequent rise in levels of prostaglandins. Evidence is mounting to suggest that the cPLA<sub>2</sub> functions in a signal-transduction pathway leading to the rapid release of arachidonic acid from membranes for the biosynthesis of eicosanoids (48-50). The Ca<sup>2+</sup>-independent PLA<sub>2</sub> may be involved in the degradation of membranes that accompanies myocardial infarction (34) and in the liberation of free arachidonate in hormone-stimulated aortic smooth muscle cells (51). For all of these reasons, most major pharmaceutical companies have ongoing programs aimed at discovering new inhibitors of PLA<sub>2</sub>s.

The analysis of PLA<sub>2</sub> inhibition is more difficult than for most enzymes because the enzymatic hydrolysis of naturally occurring, long-chain phospholipids necessarily occurs at the lipid/water interface. The present article focuses on the design and analysis of tight-binding inhibitors of PLA<sub>2</sub>s that have been carried out in the authors' laboratories. The rationale for the design of the inhibitors is discussed because it relates to the catalytic mechanism of PLA<sub>2</sub>; this is revealed by X-ray diffraction analysis of PLA<sub>2</sub>-inhibitor complexes. The problem of defining the mode of action of inhibitors of interfacial catalysis by PLA<sub>2</sub>s is also discussed, as is the development of experimental techniques for analyzing the inhibition. Finally, initial results of PLA<sub>2</sub> inhibition studies with living cells and organs are discussed as they bear on the biological functions of PLA<sub>2</sub>s.

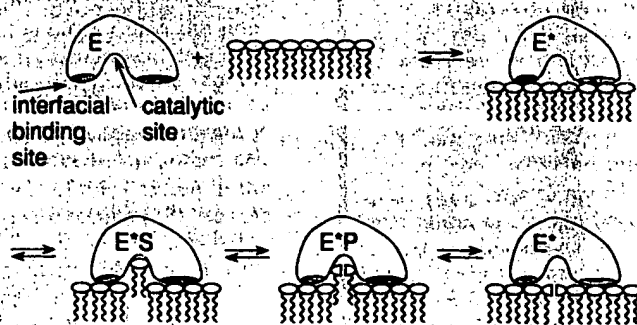
### INTERFACIAL CATALYSIS BY PLA<sub>2</sub>: THE KINETIC PARADIGM

Before discussing PLA<sub>2</sub> inhibition, it is necessary to briefly describe some basic features of interfacial catalysis by this enzyme. As naturally occurring phospholipids are essentially

insoluble in water and thus form aggregates, PLA<sub>2</sub>s must be able to bind to such aggregates in order to gain access to substrate. Two modes of interfacial catalysis can be envisioned, and are illustrated in Fig. 1. In mode 1, it is imagined that the enzyme has a surface that can anchor it to a phospholipid aggregate such as a bilayer and that the enzyme also contains a topologically distinct active site where the lipolysis reaction takes place. In this mode, the enzyme is able to bind to the bilayer without a phospholipid molecule bound in its active site; the active site binding occurs after the enzyme binds to the interface. Thus, the following enzyme species occur in mode 1: E is the enzyme in the aqueous phase, E\* is the enzyme bound to the interface but with its active site devoid of substrate and filled with solvent, E\*S is the enzyme bound to the interface and with its active site occupied by a single substrate molecule (E\*S is analogous to the well-known Michaelis complex, ES, for an enzyme operating in a homogeneous, aqueous-phase environment), and E\*P is the enzyme bound to the interface with lipolysis products (fatty acid and lysophospholipid) in its active site. Interfacial catalysis may not necessarily proceed according to mode 1. As shown in Fig. 1, mode 2 can be imagined in which the binding of the enzyme to the interface and the loading of its active site with a molecule of substrate are the same step. In this case, the enzyme species E\* does not exist.

As will become clear throughout this review, the distinction between these modes is critical for a proper understanding of most aspects of interfacial catalysis including inhibition. The distinction between these modes comes from considering the possibility of processive interfacial catalysis. In the present context, processivity means that the enzyme can remain bound to a phospholipid vesicle during the course of more than one lipolysis turnover cycle. The most extreme type of processivity is when the enzyme hydrolyzes all the substrate in the outer layer of a phospholipid vesicle without ever leaving the surface. This has been called interfacial catalysis in the "scooting mode" (52) and is schematically illustrated in Fig. 2. Scooting mode catalysis has been observed for sPLA<sub>2</sub>s from a variety of sources (venoms, pancreatic juices, and human synovial fluid) (52-54) and for the cPLA<sub>2</sub> from human U937 cells (55). The fact that processive catalysis occurs rules out mode 2 from further consideration. This is because in mode 2, the enzyme necessarily lets go of the interface when the lipolysis products are released from the active site, and the enzyme has no "memory" to enable it to return to the same phospholipid vesicle from which it was previously bound. On the other hand, in mode 1 catalysis the binding of the enzyme to the interface (E → E\*) is independent of the binding of substrate in the active site (E\* + S → E\*S). If the E to E\* step is sufficiently favorable so that the residency time of the enzyme on the interface is longer than the time it takes to complete a lipolysis reaction, processive catalysis will occur. Spectroscopic and chromatographic techniques have shown that sPLA<sub>2</sub>s bind to phospholipid vesicles and phospholipid-containing detergent micelles (2, 56-58), but the fact that this binding is independent of the binding of a substrate molecule to the active site comes only from the demonstration that the reaction can be processive (52). Also, pig pancreatic sPLA<sub>2</sub>, containing an active site-directed alkylating agent (*p*-bromo-phenacyl group attached to the catalytic histidine residue), is still able to bind to phospholipid aggregates even though substrate cannot bind to the active site because of steric clash with the inactivator (59). To the extent that this property of the inactivated enzyme applies to the native enzyme, this result is further evidence for the mode 1 nature of PLA<sub>2</sub>-catalyzed reactions.

Mode 1: Binding of E to the interface and binding of S to the catalytic site are distinct steps



Mode 2: Binding of E to the interface and binding of S to the catalytic site are the same step

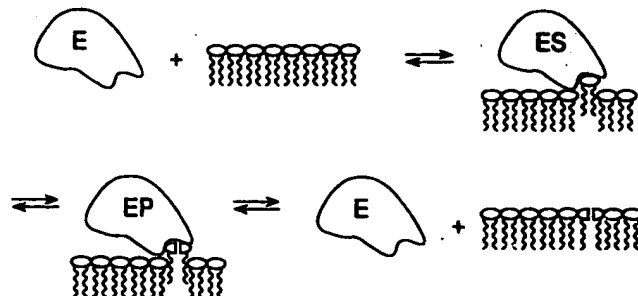
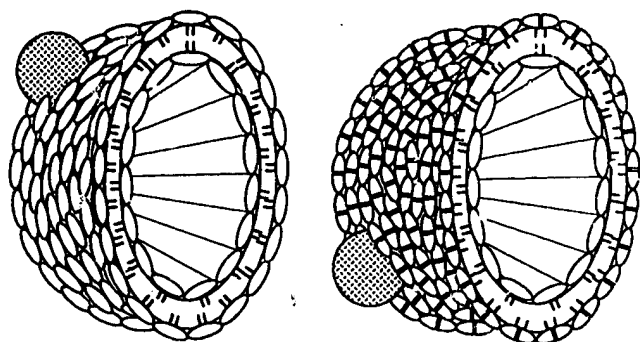


Figure 1. Two possible modes of interfacial catalysis by PLA<sub>2</sub>s.





**Figure 2.** Schematic illustration of the  $\text{PLA}_2$ -catalyzed hydrolysis of vesicles in the highly processive scooting mode. The enzyme (small sphere) remains irreversibly bound to the vesicle, and only the phospholipids in the outer leaflet of the bilayer are hydrolyzed (the vesicle is sliced in half to reveal the inner leaflet phospholipids, which are not hydrolyzed).

Direct binding studies show that  $\text{sPLA}_2$ s bind several orders of magnitude tighter to vesicles that are anionic (such as vesicles of phosphatidylmethanol or mixed-lipid vesicles of phosphatidylcholine with phosphatidic acid or phosphatidylserine) compared with vesicles that are charge neutral (such as vesicles composed only of phosphatidylcholine) (16, 52, 54, 57, 60). The enzyme hydrolyzes these anionic vesicles in the scooting mode. With vesicles composed only of phosphatidylcholine,  $\text{sPLA}_2$ s catalyze interfacial reactions in a less processive mode in which the enzyme hops from one vesicle to another (57). In this case, the reaction progress curves are complex. The processivity and reaction velocity increase as the reaction progresses because accumulation of the fatty acid reaction product gives the vesicles a net negative charge that promotes enzyme binding (61). Kinetic modeling studies suggest that the complex reaction progress curve for the action of snake venom and pancreatic  $\text{sPLA}_2$ s on large unilamellar phosphatidylcholine vesicles results from dimerization of the enzyme bound to vesicles, and that this dimer persists as the enzyme hops from one vesicle to another (62). Enzyme dimerization, if it occurs, is not obligatory for interfacial catalysis as it has been conclusively shown that monomeric  $\text{sPLA}_2$  is fully active on anionic vesicles in the scooting mode, and the enzyme does not aggregate even when multiple enzymes are present on each vesicle (53).  $\text{cPLA}_2$ s also undergo scooting mode catalysis on vesicles of phosphatidylcholine containing 5–10 mol % of reaction products (1:1 mixture of lysophospholipid and fatty acid) (55). Numerous straightforward experimental protocols have been developed to analyze for scooting behavior (55, 63), and these are not discussed in this review.

Analysis of  $\text{PLA}_2$ s in the scooting mode provides the most straightforward method for proper analysis of the kinetics of interfacial catalysis. For enzymes that operate in the aqueous phase, it is assumed in all kinetic treatments that each enzyme in the mixture "sees" the same global concentrations of substrates and products. This is not necessarily the case for interfacial enzymes. Consider interfacial catalysis under conditions in which the processivity is low. At some point, a  $\text{PLA}_2$  may leave a vesicle and bind to a new one that has not yet had a bound enzyme. This enzyme is effectively at time zero because it encounters only substrate and not products. In another instance, an enzyme may land on a vesicle that is mostly hydrolyzed; this enzyme is effectively at a long reaction time because it experiences product inhibition and substrate depletion. Thus, low processive catalysis leads to a

scrambling of time, making kinetic analyses difficult, if not impossible, to interpret. This problem is eliminated by studying  $\text{PLA}_2$ s in the scooting mode. Consider the situation in which the enzymes are irreversibly bound to the vesicles and there is sufficient excess of vesicles over enzymes so that vesicles contain at most one bound enzyme. Now each enzyme-containing vesicle will be synchronized in time; at time zero the enzyme sees a full complement of substrate and at the end of the reaction the substrate in the vesicles has been depleted. Thus, the total product formed at any time is simply the amount of product formed in an enzyme-containing vesicle multiplied by the number of such vesicles.

The scooting mode analysis gives, at first glance, an artificial impression of interfacial catalysis, as under physiological conditions the  $\text{PLA}_2$  may or may not be operating in a highly processive mode. However, for the reasons given above, the scooting mode analysis allows the kinetic and equilibrium constants that describe the interfacial catalysis to be measured, and these constants will apply to the action of the enzyme in the interface even if the enzyme is coming on and off the membrane. Under the conditions of low processivity, the interfacial kinetic and equilibrium parameters will be difficult to obtain as they are convoluted with the equilibrium constants that describe the binding of the enzyme to the interface; the latter can be measured directly by spectroscopic techniques (63). The problem is further complicated by the fact that the affinity of  $\text{PLA}_2$ s for the interface depends on the composition of the interface and thus on the fraction of substrate that has been converted to products. All these points have been discussed in detail elsewhere (57, 63, 64).

#### THE STRUCTURAL BASIS FOR SCOOTING MODE BEHAVIOR OF $\text{PLA}_2$

As stated in the introduction, X-ray structures are available for several  $\text{sPLA}_2$ s and, in some cases, with short-chain phospholipid analogs bound in their active sites. The structures reveal that the active site consists of a slot that traverses the enzyme into which a single phospholipid molecule can bind and interact with the catalytic residues (Fig. 3). This slot is clearly distinct from the surface of the enzyme that is thought to contact the bilayer (60). These structures and the kinetic results that demonstrate that the  $\text{E} \rightarrow \text{E}^*$  step is distinct from the  $\text{E}^* + \text{S} \rightarrow \text{E}^*\text{S}$  step (52, 60) suggest that a phospholipid molecule must diffuse about 15 Å from the plane of the bilayer into the enzyme bound to the interface in order to reach the catalytic machinery.

These high-resolution structures have also provided a wealth of information about the way in which the enzyme utilizes a histidine residue and a  $\text{Ca}^{2+}$  ion to help bind the substrate in the active site and to stabilize the substrate-derived transition state (8, 10).

#### TRUE COMPETITIVE INHIBITORS OF INTERFACIAL CATALYSIS BY $\text{PLA}_2$

The scooting mode analysis is also important for the proper analysis of reversible-binding, competitive inhibitors of  $\text{PLA}_2$ . History has shown that it is problematic to study reversible inhibitors of  $\text{PLA}_2$  under conditions in which the enzyme is weakly bound to the interface (low processivity). This is because it is difficult to sort out whether an inhibitor is working by competing with the substrate for binding to the active site ( $\text{E}^* + \text{S} \rightarrow \text{E}^*\text{S}$ ) or by altering the physical nature

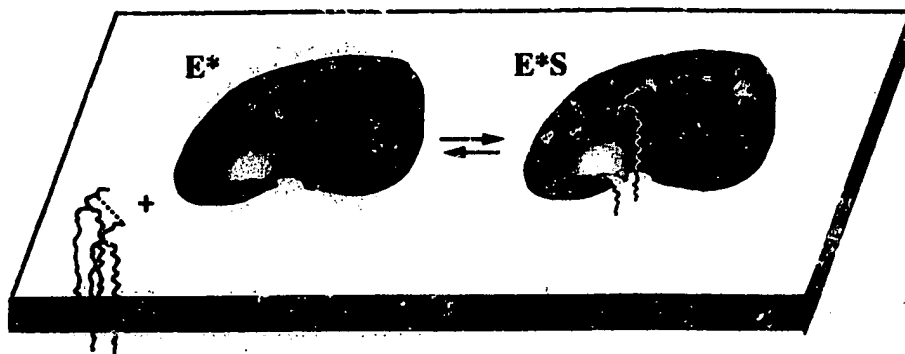


Figure 3. Movement of the substrate from the plane of the bilayer into the catalytic slot of the vesicle-bound sPLA<sub>2</sub>. In the bottom-left corner are two phospholipids that interact via their polar head groups (dashed line). The polar head group of the phospholipid in the catalytic slot of the sPLA<sub>2</sub> is shown interacting with an amino acid side chain (dashed line); such an interaction occurs with the cobra venom sPLA<sub>2</sub> but not with the enzymes from pancreas, human synovial fluid, and bee venom.

of the interface so that more of the enzyme desorbs from the surface ( $E^* \rightarrow E$ ). Many previously reported PLA<sub>2</sub> inhibitors have been shown to partition into vesicles and to decrease the fraction of enzyme bound to vesicles (65-67). The characteristics of such nonspecific and therefore uninteresting PLA<sub>2</sub> inhibitors are as follows: 1) inhibition is seen when the enzyme is weakly bound to vesicles but not when the enzyme is tightly bound to vesicles (scooting mode); 2) stereoisomers of inhibitors with chiral centers display similar inhibition potencies; 3) large amounts of inhibitors (corresponding to more than a few mole % inhibitor in the vesicle) are needed to produce the inhibition. The last point is often overlooked. Many PLA<sub>2</sub> assays are carried out radio-metrically with small amounts of phospholipid substrate (micromolar), and thus even nanomolar amounts of inhibitor can have a profound effect on the physical nature of the vesicle surface especially for inhibitors that partition favorably into vesicles (contrast this with the situation of enzymes that operate in the aqueous phase where the inhibitor and substrate do not form a coaggregate).

All these problems are avoided by analyzing putative PLA<sub>2</sub> inhibitors in the scooting mode. The presence of additives in anionic vesicles up to 10-20 mol % do not affect the kinetics of interfacial catalysis as the enzyme remains tightly bound to the interface (65). Instead, agents that bind specifically to the active site of the enzyme, the true competitive inhibitors of interfacial catalysis, are detected. The scooting mode analysis has been useful for conducting large-scale screens of compound banks for PLA<sub>2</sub> inhibitors as is routinely done at many pharmaceutical companies (63). Such an analysis does not detect false positive lead compounds that are often encountered in inhibitor screens. By using an assay in which the enzyme is tightly bound to the vesicle, possible inhibitors that bind specifically and tightly to the enzyme in the aqueous phase and prevent the enzyme from binding to the interface may be missed; this is a perfectly valid and useful mode of PLA<sub>2</sub> inhibition. However, the molecular surface of the PLA<sub>2</sub> that contacts the interface is large, and the membrane binding involves the interaction of several amino acids with the interface; thus, it is unlikely that an inhibitor will be able to mask a fraction of this surface sufficient to prevent the  $E \rightarrow E^*$  step. Inhibitors that bind in the active site of the enzyme in solution but do not prevent the  $E \rightarrow E^*$  step can also be imagined, but if such compounds are driven off the enzyme concomitant with the  $E \rightarrow E^*$  step, the inhibitor will be useless. In other words,

one is forced to search for inhibitors that may or may not bind to E but must bind to E\*. Thus, there is little merit in using inhibitor screening strategies that employ E acting on a water-soluble substrate. Other aspects of these arguments have been discussed previously (69).

The equation that describes the inhibition (Eq. 1) is a straightforward extension of the standard steady-state equation that describes competitive inhibition of enzymes that operate in aqueous solution:

$$v_0/v_0^0 = 1/(1 + 1/K_I^*)/(1 + 1/K_M^*)[(X_1)/(1 - X_1)] \quad (1)$$

Here,  $v_0$  and  $v_0^0$  are the initial reaction velocities for the action of the enzyme in the scooting mode in the absence or presence of a competitive inhibitor, respectively.  $K_M^*$  is the interfacial Michaelis constant [just as  $K_M$  is the concentration (mol/l) of substrate in the aqueous phase that gives a reaction velocity of  $V_{max}/2$ ,  $K_M^*$  is the mole fraction of substrate in the vesicle interface that gives  $V_{max}/2$ ].  $K_I^*$  is the interfacial dissociation constant for the  $E^*I$  complex (the mole fraction of I in the interface that gives  $E^*/E^*I = 1$ ). Thus, the degree of inhibition depends on the affinity of the inhibitor and substrate to the enzyme ( $K_I^*$  and  $K_M^*$ ) and the mole fraction of inhibition in the vesicle ( $X_1$ ) [if the amount of inhibitor in the interface is significantly less than the amount of substrate in the vesicle, as is the case for tight-binding inhibitors, the  $(1 - X_1)$  term in Eq. 1 can be ignored]. True competitive inhibitors of PLA<sub>2</sub> lead to kinetics that are consistent with Eq. 1 (65, 66).

Phospholipids dispersed in detergent micelles have also been used as PLA<sub>2</sub> substrates and in the analysis of inhibitors (see, for example, refs 70-73). In these systems, tight-binding competitive inhibitors are revealed, and the degree of inhibition depends on the stereochemistry of chiral inhibitors, as is the case with vesicle assays. Some differences between mixed micelles and vesicle assays of inhibition may be noted. The binding of sPLA<sub>2</sub> to mixed micelles is typically of lower affinity than to anionic vesicles, and thus the mole fraction of inhibitor in mixed micelles should be kept low so that one avoids the problem of a shift in the E to E\* equilibrium as previously discussed. It is becoming apparent that reaction velocities for sPLA<sub>2</sub>-catalyzed lipolysis of mixed micelles is limited by the rate of substrate replenishment in enzyme-containing aggregates (74, 75). This stems from the

fact that mixed micelles, in contrast to vesicles, contain only a few substrate molecules, and thus significant substrate depletion occurs in the millisecond time scale. Interaggregate exchange of enzyme and/or phospholipid is required to maintain the reaction progress. Thus, it is possible to imagine an inhibitor that functions not by binding directly to  $E^*$  but by slowing the rate of interaggregate exchange of substrate and/or enzyme. It may or may not be a valid assumption that such a concern will be avoided if the mole fraction of inhibitor in the aggregate is kept low. Finally, the process or rate-limiting substrate replenishment may distort the value of the mole fraction of inhibitor needed to cause 50% inhibition. This is because the local concentration of substrate in enzyme-containing mixed micelles will be lower than that in the bulk of the micelles that do not have bound enzyme. Thus, the mole fraction of inhibitor needed to cause a 50% reduction will be lower than that predicted by the true  $K_I^*$  value, because the inhibitor is competing with an artificially low amount of substrate present in enzyme-containing mixed micelles. This inequality is not possible if the rate of substrate replenishment is fast on the lipolysis turnover time scale so that the local and global steady-state substrate concentrations are equivalent. This situation is not a serious problem as values of  $K_I^*$  determined kinetically and under equilibrium conditions are correlated as expected (76).

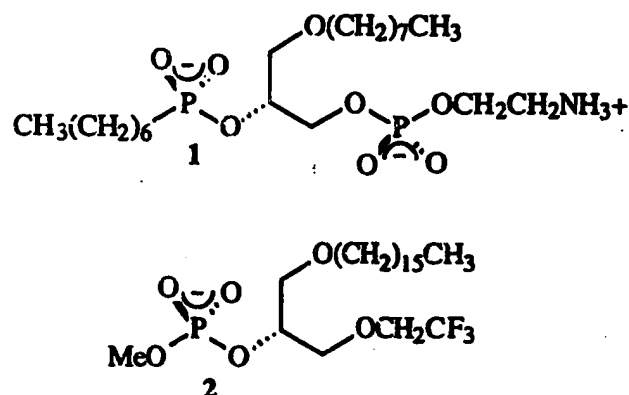
### NEUTRAL DILUENTS

Because binding of the enzyme to the interface is independent of the binding of a molecule in the interface to the active site of the enzyme, it should be possible to find an amphiphile that forms an aggregate to which  $PLA_2$  can bind but has no affinity for the active site of the enzyme and so  $E^*$  has its active site filled with water. Such a compound has been termed a neutral diluent because it can be added to a vesicle of substrate phospholipid and thereby dilute the surface concentration of substrate that the bound enzyme "sees," and at the same time the neutral diluent does not inhibit the enzyme by binding to its catalytic site. The amphiphiles 2-hexadecyl-*sn*-glycero-3-phosphocholine and hexadecylphosphocholine are neutral diluents for pig pancreatic and human synovial  $sPLA_2$ s, respectively (54, 65). This has been established by the demonstration that active site-directed alkylating agents (*p*-substituted phenacyl bromides) modify the catalytic histidine residue of the enzyme in the aqueous phase or bound to an aggregate of neutral diluent at the same rate, which shows that the active site is not protected from alkylation by the binding of the amphiphile. The use of a neutral diluent in combination with an active site-directed alkylating agent provides an independent measure of  $K_I^*$  values for  $PLA_2$  inhibitors. This is because when a competitive inhibitor of  $PLA_2$  is present in an aggregate of neutral diluent at mole fraction =  $K_I^*$ , half of the bound enzyme will be in the  $E^*I$  form and the half-time for alkylation of the enzyme will be twice its value obtained in the absence of the inhibitor (65).

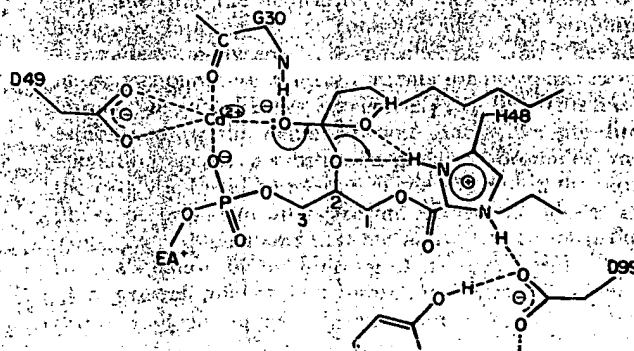
### STRUCTURAL CLASSES OF REVERSIBLE, TIGHT-BINDING COMPETITIVE INHIBITORS OF $PLA_2$

Phospholipid analogs that have a phosphonate or phosphate in place of the ester at the *sn*-2 position of phospholipids

(compounds 1 and 2 (below) are tight-binding inhibitors of  $sPLA_2$ s but not  $cPLA_2$ s (54, 77-79). These substrate analogs are patterned after the putative tetrahedral intermediate that



forms when a water molecule attacks the substrate *sn*-2 ester; they bind to  $E^*$  about three orders of magnitude tighter than substrates. The X-ray structures of  $sPLA_2$ s containing bound compound 1 suggest how the lipolysis reaction is catalyzed (8, 15) (Fig. 4). One nonbridging oxygen of the phosphonate is hydrogen bonded to the protonated imidazole ring of the active site histidine, and the other is directly liganded to the calcium ion. This suggests that the mechanism of substrate hydrolysis involves the attack of a water molecule onto the ester carbonyl carbon assisted by general base catalysis by histidine. The oxyanion, which is derived from the carbonyl oxygen, is stabilized by binding to calcium. The calcium also helps to anchor the substrate in the



**Figure 4.** The structural details of the interaction of the putative tetrahedral intermediate with the catalytic residues of  $sPLA_2$ s as revealed by X-ray diffraction. The oxyanion of the tetrahedral intermediate interacts with calcium and the amide NH of glycine-30 (pancreatic  $sPLA_2$  residue numbering). The hydroxyl group of the tetrahedral intermediate is hydrogen bonded to the protonated histidine; this amino acid likely protonates the *sn*-2-oxygen as the lysophospholipid is formed. The histidine is also hydrogen bonded to aspartate-99, which receives a hydrogen bond from a tyrosine residue. The *sn*-3-phosphate is liganded to the calcium; the latter is held to the enzyme by aspartate-49 and glycine-30, as shown, as well as two additional carbonyl oxygen ligands from the peptide backbone.

active site by its interaction with one of the nonbridging oxygens of the *sn*-3 phosphate; the other nonbridging oxygen of this phosphate is hydrogen bonded to either a tyrosine (pancreatic and cobra venom sPLA<sub>2</sub>s), a threonine (bee venom sPLA<sub>2</sub>), or a lysine (human synovial fluid sPLA<sub>2</sub>).

Compound 2 is a member of a novel family of phospholipid analogs that are tight-binding inhibitors of sPLA<sub>2</sub>s (79). These compounds have an *sn*-2 phosphate but lack one at the *sn*-3 position. Apparently the polar head group of a phospholipid is not a critical element for its interaction with the active site of sPLA<sub>2</sub>s even though it forms interactions with the active site as described in the previous section. This is also apparent from the following additional inhibitor studies. Analogs of phospholipids in which the *sn*-2 ester is replaced by an amide (-CONH-) are tight-binding inhibitors of sPLA<sub>2</sub>s (70, 71, 80). The amide hydrogen is hydrogen bonded to the nonprotonated imidazole ring of the active site histidine, and the *sn*-3 phosphate is liganded to the calcium (10, 72). Surprisingly, simple fatty acid amides (RCONH<sub>2</sub>) that lack the entire glycerol backbone bind to sPLA<sub>2</sub>s with affinities approaching those of phospholipid amide analogues (discussed below) (68).

None of the sPLA<sub>2</sub> inhibitors described cause detectable inhibition of cPLA<sub>2</sub>. However, an analog of arachidonic acid in which the COOH group is replaced with a trifluoromethyl ketone (COCF<sub>3</sub>) is a tight-binding, reversible inhibitor of cPLA<sub>2</sub>; this compound (AACOCF<sub>3</sub>) is more than 1,000-fold less potent as an inhibitor of sPLA<sub>2</sub>s (73). NMR studies suggest that AACOCF<sub>3</sub> forms a hemiketal adduct with an active site nucleophile [Enz-X-C(OH)(CF<sub>3</sub>)-AA] (81). Such an adduct may resemble the tetrahedral intermediate in route to a putative acyl enzyme, but further studies will be required to fully understand the mechanism of catalysis by cPLA<sub>2</sub>. Analogs of AACOCF<sub>3</sub> in which the COCF<sub>3</sub> group is replaced by COCH<sub>3</sub>, CH(OH)CF<sub>3</sub>, or CONH<sub>2</sub> are not cPLA<sub>2</sub> inhibitors (73). A possible reason for the ability of inhibitors that lack a phospholipid-like polar head group to bind tightly to sPLA<sub>2</sub> and cPLA<sub>2</sub> is discussed in the next section.

#### DIFFERENTIAL INTERACTIONS OF INHIBITORS WITH THE PHOSPHOLIPID INTERFACE AS A CONTROLLING FACTOR FOR POTENCIES OF PLA<sub>2</sub> INHIBITORS

In considering the factors that determine the magnitude of enzyme-inhibitor dissociation constants, one must examine not only the interactions that the inhibitor engages in with the enzyme but also the interactions that the inhibitor forms with its environment when it is not bound to the enzyme. For enzymes in the aqueous phase, experimental and theoretical studies have shown that differential solvation of inhibitors can dictate their relative affinities for enzymes (see, for example, ref 82). Thus, compound A may be a better inhibitor than B not because it forms stronger interactions with the enzyme than B but because as a free species it is less solvated than free B. The same is true for inhibitors of PLA<sub>2</sub>, except that the free inhibitor is now in a lipid vesicle instead of in aqueous solution. Thus, differential interactions of inhibitors with adjacent phospholipid molecule in the vesicle can control, in part, the relative affinities of inhibitors for PLA<sub>2</sub> bound to the vesicle surface.

A systematic study was carried out on the effect of varying the structure of the polar head group of *sn*-2 phosphonate containing phospholipid analogs on the binding to sPLA<sub>2</sub>s

(83). The inhibitor potencies were obtained according to Eq. 1 and for the action of a number of sPLA<sub>2</sub>s on vesicles of 1,2-dimyristoyl-*sn*-glycero-3-phosphomethanol in the scooting mode. When comparing inhibitors of PLA<sub>2</sub>, it is important to determine whether all of the inhibitors are fully partitioned into the vesicle, otherwise the differences in inhibitor potencies may reflect differences in surface concentrations of inhibitor due to differential partitioning rather than differences in intrinsic affinities of the inhibitors for the enzyme ( $K_1^*$ ). The scooting mode analysis provides a simple way to assess inhibitor partitioning. The degree of inhibition by a fixed mole fraction of inhibitor in the vesicle is determined at different total reaction volumes. Because the enzyme is tightly bound to the vesicles, the reaction velocity in the absence of inhibitor is independent of the reaction volume. If an inhibitor is present and if it is only partially partitioned into the vesicle phase, the degree of inhibition will decrease as the reaction volume is increased because the latter change will lead to a decrease in the mole fraction of inhibitor in the interface as more and more of it partitions into the aqueous phase. By this criterion, it has been shown that the differences in inhibitor potencies reflect differences in their values of  $K_1^*$  (83). A more direct method for determining inhibitor/vesicle partition coefficients has also been reported (84).

An interesting trend was discovered from this analysis. In the case of the sPLA<sub>2</sub>s from porcine and bovine pancreas, bee venom, and human synovial fluid, replacement of the ammonium of the polar head group compound 1 with hydrogen increases the potency by 5- to 10-fold. In contrast, such a change produces no effect on the  $K_1^*$  measured with the cobra venom sPLA<sub>2</sub> (83). X-ray structural studies of sPLA<sub>2</sub> inhibitor complexes show that for the enzymes from pancreas, bee venom, and human synovial fluid, the ammonium of the phosphoethanolamine polar head group of the bound phospholipid analog does not directly contact the enzyme. Thus, the increase in inhibitor potency when the ammonium is removed must be because the ammonium group of the free inhibitor forms favorable interactions with the head group phosphate of an adjacent phospholipid in the vesicle; such an interaction is not possible with the analog lacking the ammonium group (Fig. 3). By analyzing the relative binding energies of a series of inhibitors with modified polar head groups to certain sPLA<sub>2</sub>s, it is possible to determine the relative energetics of intermolecular polar head group interactions of phospholipids in bilayered vesicles. In the case of the cobra venom enzyme, removing this ammonium group breaks the interaction of the inhibitor with the enzyme (Fig. 3) and with the bilayer; thus, such a change is expected to produce little, if any, effect on the  $K_1^*$  values, as is observed.

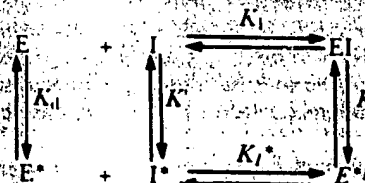
#### E\* HAS INTRINSICALLY HIGHER AFFINITY FOR INHIBITORS COMPARED WITH E

The binding of phospholipid analogs containing short hydrocarbon chains to PLA<sub>2</sub> in the aqueous phase can be studied because such compounds are water soluble as solitary monomers below their critical micelle concentrations or solubility limit. The apparent affinity of phosphonate phospholipid analog 1 for cobra venom sPLA<sub>2</sub> in the aqueous phase increases enormously in the presence of the non-hydrolyzable short-chain substrate analog 1,2-dihexyl-*sn*-glycero-3-phosphocholine (85). In this study, the binding of radiolabeled inhibitor 1 to cobra venom sPLA<sub>2</sub> was measured by equilibrium dialysis, and the  $K_1$  for the enzyme-inhibitor complex is about 50  $\mu$ M. Thus, in the aqueous



phase, compound 1 acts as a rather weak-binding inhibitor. The addition of 1,2-dihexyl-*sn*-glycero-3-phosphocholine to the water-soluble cobra venom enzyme leads to the formation of a protein/lipid microaggregate that can be detected by gel filtration or analytical ultracentrifugation (86-88). Aggregation is also seen with the pancreatic sPLA<sub>2</sub> in the presence of short-chain anionic amphiphiles (89, 90). These aggregates presumably form by the ability of the interfacial recognition surface of the PLA<sub>2</sub> to recruit short-chain, water-soluble phospholipids in a process that leads to the segregation of such complexes to form a protein-lipid microaggregate. Thus, in the presence of 1,2-dihexyl-*sn*-glycero-3-phosphocholine, the cobra venom PLA<sub>2</sub> may have the properties more of E\* than of E. The fact that the inhibitor binds more tightly to the enzyme in the presence of this microaggregation suggests that E\* has intrinsically higher affinity for E\* than does E.

These results are only qualitative, and they have sparked the design of additional and more rigorous experiments that address the intrinsic affinity of E vs. E\* for active site ligands (84). One problem that immediately arises is that it is not possible to directly compare the values of K<sub>i</sub> (for the process EI <=> E + I) to K<sub>i</sub>\* (for the process (E\*I <=> E\* + I\*), as the former has units of molarity and the latter has units of mole fraction. Instead it is most useful to consider these equilibrium processes in terms of Scheme II shown below.



Scheme II

In this scheme all the underlying equilibria are considered for the interaction of the inhibitor with the enzyme both in the aqueous phase and at the interface. Thus, in addition to the dissociation constants K<sub>i</sub> and K<sub>i</sub>\* discussed previously, the constants K<sub>d</sub> and K<sub>d</sub>\* are the equilibrium constants for the dissociation of E\* and E\*I, respectively, from the interface of the neutral diluent into the aqueous phase. Thus, K<sub>d</sub> = [neutral diluent][E]/[E\*] and K<sub>d</sub>\* = [neutral diluent][E\*I]/[E\*I] where quantities in brackets are in units of moles per unit volume. The remaining constant, K', is the equilibrium constant for the dissociation of the inhibitor from the interface into the aqueous phase (K' = [neutral diluent][I]/[I\*]).

Equation 2 is a thermodynamic relation (detailed-balance condition) derived from Scheme II.

$$K_d/K_d^* = K_i/(K_i^* K') = [E][E^*I]/[E^*][EI] \quad (2)$$

According to this equation, if K<sub>d</sub>/K<sub>d</sub>\* > 1, [E\*I]/[E\*] > [EI]/[E], and thus regardless of the total amount of inhibitor in the system, the fraction of enzyme in the interface that is bound to inhibitor is higher than that for the enzyme in the

aqueous phase. Using the thermodynamic cycle (Scheme II) in this way provides a direct evaluation of the intrinsic inhibitor affinity and automatically accounts for differences in local concentrations of inhibitor in the aqueous phase and at the interface.

For all true competitive inhibitors of the pig pancreatic sPLA<sub>2</sub> studied, it is found that K<sub>d</sub>/K<sub>d</sub>\* is approximately 50 (84). This strengthening of the enzyme-interface binding could come about in three different ways: 1) It could be due to a residual inhibitor-interface interaction where part of the inhibitor remains attached to the interface even when the inhibitor has entered the catalytic site; 2) The interface could somehow be influenced by the inhibitor moving into the catalytic site so that it makes better contact with the enzyme; 3) The enzyme may undergo a conformation change when the inhibitor binds so that it makes better contact with the interface. The first two possibilities would suggest that the length of the alkyl chains of the inhibitors would influence the strength of the effect; surprisingly, this is not the case (84). Thus, the most likely effect is an allosteric modulation of the enzyme that occurs when the inhibitor binds in the catalytic site.

## PLA<sub>2</sub> INHIBITION STUDIES IN LIVING CELLS

Now that some PLA<sub>2</sub>s in cells have been identified and true competitive inhibitors that are enzyme selective have been developed, it should be possible to take a pharmacological approach toward determining the roles that these enzymes play in biological processes. Of course, such an approach does not provide unequivocal data because unidentified PLA<sub>2</sub>s may exist in cells, and the full spectrum of enzyme selectivity of any given inhibitor is not known a priori.

The presence of AACOCF<sub>3</sub> at concentrations of 0-20 μM blocks essentially all the arachidonic acid that is liberated in thrombin-stimulated platelets (91, 92), in calcium ionophore-stimulated human monocytic cells (U937) (92), and in interleukin 1α-stimulated mesangial cells (93). However, such effects are not seen when the cells are treated with analogs of AACOCF<sub>3</sub> which have minor structural alterations that obliterate their abilities to inhibit cPLA<sub>2</sub> (see above). Such results argue strongly for the role of the cPLA<sub>2</sub> in liberating arachidonate for the biosynthesis of eicosanoids. Furthermore, they allay concerns about using membrane-residing inhibitors to inhibit the action of cPLA<sub>2</sub> in membranes.

Phospholipid analog 2 protects isolated rat lungs from ischemia reperfusion injury when it is delivered as a suspension in liposomes, whereas micelles of this compound added directly to the perfusion medium are ineffective (36). Compound 2 also inhibits the degradation of surfactant phosphatidylcholine by rat lung (94). Biochemical studies suggest that the primary inhibitory effect of this compound is on a calcium-independent PLA<sub>2</sub> of lysosomal origin with optimum activity at acidic pH. This activity is probably involved in the turnover of phospholipids of lung surfactant. This enzyme not only hydrolyzes normal phospholipids but is also responsible for the selective release of conjugated dienes and thiobarbituric acid reactive substances; however, the phospholipid cleavage leading to the formation of thromboxane is not affected.

A substituted bromoenoyl lactone is a specific inhibitor of the calcium-independent PLA<sub>2</sub> from myocardium, and this compound blocks arachidonate liberation in hormonally stimulated aortic smooth muscle cells (51). [51]

This work was supported by grants from the National Institutes of Health (HD36235 and a Research Career Development Award, GM562, to M.H.G. and GM29703 to M.K.J.), grants from Sterling Winthrop Inc. to M.H.G. and M.K.J., and a Sloan Fellowship to M.H.G. The authors are grateful to Drs. Dennis Murphy (Sterling-Winthrop) and Ian Street (Merck Frosst) for helpful discussions.

## REFERENCES

- Waite, M. (1987) *The phospholipases*. Plenum, New York
- Verheij, H. M., Slotboom, A. J., and de Haas, G. H. (1981) Pancreatic phospholipase A<sub>2</sub>: a model for membrane-bound enzymes? *Rev. Physiol. Biochem. Pharmacol.* **91**, 91-203
- Dennis, E. A. (1983) Phospholipases. *The Enzymes* **16**, 307-353
- Kramer, R. M., Jakubowski, J. A., and Deykin, D. (1988) Hydrolysis of 1-alkyl-2-arachidonoyl-sn-glycero-3-phosphocholine, a common precursor of platelet-activating factor and eicosanoids, by human platelet phospholipase A<sub>2</sub>. *Biochim. Biophys. Acta* **959**, 269-279
- Seilhamer, J. J., Pruzanski, W., Vadas, P., Plant, S., Miller, J. A., Kloss, J., and Johnson, L. K. (1989) Cloning and recombinant expression of phospholipase A<sub>2</sub> present in rheumatoid arthritic synovial fluid. *J. Biol. Chem.* **264**, 5335-5338
- Yu, B.-Z., Berg, O. G., and Jain, M. K. (1993) The divalent cation is obligatory for the binding of ligands to the catalytic site of secretory phospholipases A<sub>2</sub>. *Biochemistry* **32**, 6485-6492
- van den Bergh, C. J., Slotboom, A. J., Verheij, H. M., and de Haas, G. H. (1989) The role of Asp-49 and other conserved amino acids in phospholipase A<sub>2</sub> and their importance for enzymatic activity. *J. Cell Biochem.* **39**, 379-390
- Scott, D. L., White, S. P., Orwinowski, Z., Yuan, W., Gelb, M. H., and Sigler, P. B. (1990) Interfacial catalysis: the mechanism of phospholipase A<sub>2</sub>. *Science* **250**, 1541-1546
- Wéry, J.-P., Schevitz, R. W., Clawson, D. K., Bobbitt, J. L., Dow, E. R., Gamboa, G. C., Goodson, J., Hermann, R. B., Kramer, R. M., McClure, D. B., Mihelich, E. D., Putnam, J. E., Sharp, J. D., Stark, D. H., Teater, C., Warrick, M. W., and Jones, N. D. (1991) Structure of recombinant human rheumatoid arthritic synovial fluid phospholipase A<sub>2</sub> at 2.2-angstroms resolution. *Nature (London)* **352**, 79-82
- Thunnissen, M. M. G. M., Ab, E., Kalk, K. H., Drenth, J., Dijkstra, B. W., Kuipers, O. P., Dijkman, R., de Haas, G. H., and Verheij, H. M. (1990) X-ray structure of phospholipase A<sub>2</sub> complexed with a substrate-derived inhibitor. *Nature (London)* **347**, 689-691
- Fremont, D. H., Anderson, D. H., Wilson, I. A., Dennis, E. A., and Xuong, N. H. (1993) Crystal structure of phospholipase A<sub>2</sub> from Indian cobra reveals a trimeric association. *Proc. Natl. Acad. Sci. USA* **90**, 342-346
- Holland, D. R., Clancy, L. L., Muchmore, S. W., Ryde, T. J., Einspahr, H. M., Finzel, B. C., Henrikson, R. L., and Warrenpaugh, K. D. (1990) The crystal structure of a lysine 49 phospholipase A<sub>2</sub> from the venom of the cottonmouth snake at 2.0-Å resolution. *J. Biol. Chem.* **265**, 17649-17656
- Westerlund, B., Nordlund, P., Uhlin, U., Eaker, D., and Eklund, H. (1992) The three-dimensional structure of notexin, a presynaptic neurotoxic phospholipase A<sub>2</sub> at 2.0-Å resolution. *FEBS Lett.* **301**, 159-164
- Tomoo, K., Ohishi, H., Doi, M., Ishida, T., Inoue, M., Ikeda, K., Hata, Y., and Samejima, Y. (1992) Structure of the acidic phospholipase A<sub>2</sub> for the venom of *Akistrodon halys blomhoffi* at 2.8-Å resolution. *Biochem. Biophys. Res. Commun.* **184**, 137-143
- Scott, D. L., White, S. P., Browning, J. L., Rosa, J. J., and Gelb, M. H. (1991) Structures of free and inhibited human secretory phospholipase A<sub>2</sub> from inflammatory exudate. *Science* **254**, 1007-1010
- Ghomashchi, F., Yu, B.-Z., Berg, O., Jain, M. K., and Gelb, M. H. (1991) Interfacial catalysis by phospholipase A<sub>2</sub>: substrate specificity in vesicles. *Biochemistry* **30**, 7318-7329
- Kim, D. K., Kudo, I., and Inoue, K. (1991) Purification and characterization of rabbit platelet cytosolic phospholipase A<sub>2</sub>. *Biochim. Biophys. Acta* **1083**, 80-88
- Takayama, K., Kudo, I., Kim, D. K., Nagata, K., Nozawa, Y., and Inoue, K. (1991) Purification and characterization of human platelet phospholipase A<sub>2</sub> which preferentially hydrolyses an arachidonoyl residue. *FEBS Lett.* **282**, 326-330
- Kim, D. K., Suh, P. G., and Ryu, S. H. (1991) Purification and some properties of a phospholipase A<sub>2</sub> from bovine platelets. *Biochem. Biophys. Res. Commun.* **174**, 189-196
- Leslie, C. C., Voelker, D. R., Channon, J. Y., Wall, M. M., and Zelarny, P. T. (1988) Properties and purification of an arachidonoyl-hydrolyzing phospholipase A<sub>2</sub> from a macrophage cell line RAW 264.7. *Biochim. Biophys. Acta* **963**, 476-492
- Wijkander, J., and Sundler, R. (1989) A phospholipase A<sub>2</sub> hydrolyzing arachidonyl-phospholipids in mouse peritoneal macrophages. *FEBS Lett.* **244**, 51-56
- Wijkander, J., and Sundler, R. (1991) An 100-kDa arachidonate-mobilizing PLA<sub>2</sub> in mouse spleen and the macrophage cell line J774. *Eur. J. Biochem.* **2**, 873-880
- Clark, J. D., Lin, L.-L., Kriz, R. W., Ramesha, C. S., Sultzman, L. A., Lin, A. Y., Milona, N., and Knopf, J. L. (1991) A novel arachidonic acid-selective cytosolic PLA<sub>2</sub> contains a Ca<sup>2+</sup>-dependent translocation domain with homology to PKC and GAP. *Cell* **65**, 1043-1051
- Diez, E., and Mong, S. (1990) Purification of a phospholipase A<sub>2</sub> from human monocytic leukemia U-937 cells. *J. Biol. Chem.* **265**, 14654-14661
- Kramer, R. M., Roberts, E. F., Manetta, J., and Putnam, J. E. (1991) The Ca<sup>2+</sup>-sensitive cytosolic phospholipase A<sub>2</sub> is a 100-kDa protein in human monoblast U937 cells. *J. Biol. Chem.* **266**, 5268-5272
- Rehfeldt, W., Hass, R., and Goppelt-Strube, M. (1991) Characterization of phospholipase A<sub>2</sub> in monocytic cell lines. *Biochem. J.* **276**, 631-636
- Gronich, J. H., Bonventre, J. V., and Nemenoff, R. A. (1990) Purification of a high-molecular mass form of phospholipase A<sub>2</sub> from rat kidney activated at physiological calcium concentrations. *Biochem. J.* **271**, 37-43
- Channon, J. Y., and Leslie, C. C. (1990) A calcium-dependent mechanism for association of a soluble arachidonoyl-hydrolyzing phospholipase A<sub>2</sub> with membrane in the macrophage cell line RAQ 624.7. *J. Biol. Chem.* **265**, 5409-5413
- Krause, H., Dieter, P., Schulze-Specking, A., Ballhorn, A., and Decker, K. (1991) Ca<sup>2+</sup>-induced reversible translocation of phospholipase A<sub>2</sub> between the cytosol and the membrane fraction of rat liver macrophages. *Eur. J. Biochem.* **199**, 355-359
- Yoshihara, Y., and Watanabe, Y. (1990) Translocation of phospholipase A<sub>2</sub> from cytosol to membranes in rat brain induced by calcium ions. *Biochem. Biophys. Res. Commun.* **170**, 484-490
- Wijkander, J., and Sundler, R. (1992) Macrophage arachidonate-mobilizing phospholipase A<sub>2</sub>: role of Ca<sup>2+</sup> for membrane binding but not for catalytic activity. *Biochem. Biophys. Res. Commun.* **184**, 118-124
- Ghomashchi, F., Schüttel, S., Jain, M. K., and Gelb, M. H. (1992) Kinetic analysis of a high-molecular weight phospholipase A<sub>2</sub> from rat kidney: divalent metal-dependent 'trapping' of enzyme on product-containing vesicles. *Biochemistry* **31**, 3814-3824
- Sharp, J. D., White, D. L., Chiou, X. G., Goodson, T., Gamboa, G. C., McClure, D., Burgett, S., Hoskins, J., Skatrud, P. L., Sportsman, J. R., Becker, G. W., Kang, L. H., Roberts, E. F., and Kramer, R. M. (1991) Molecular cloning and expression of human Ca<sup>2+</sup>-sensitive cytosolic phospholipase A<sub>2</sub>. *J. Biol. Chem.* **266**, 14850-14853
- Gross, R. W. (1992) Myocardial phospholipases A<sub>2</sub> and their membrane substrates. *Trends Cardiovasc. Med.* **2**, 115-121
- Hirashima, Y., Farooqui, A. A., Mills, J. S., and Horrocks, L. A. (1992) Identification and purification of a calcium-independent phospholipase A<sub>2</sub> from bovine brain cytosol. *J. Neurochem.* **59**, 708-714
- al-Mehdi, A. B., Dodia, C., Jain, M. K., and Fisher, A. B. (1993) A phospholipase A<sub>2</sub> inhibitor decreases generation of thiobarbituric acid reactive substance during lung ischemia-reperfusion. *Biochim. Biophys. Acta* **1167**, 56-62
- Glaser, K. B., Mobilio, D., Chang, J. Y., and Senko, N. (1993) Phospholipase A<sub>2</sub> enzymes: regulation and inhibition. *Trends Pharmacol. Sci.* **14**, 92-98
- Bomlaski, J. S., Lawton, P., and Browning, J. L. (1991) Human extracellular recombinant PLA<sub>2</sub> induces an inflammatory response in rabbit joints. *J. Immunol.* **146**, 3904-3910
- Elsbach, P., and Weiss, J. (1993) The bactericidal/permeability-increasing protein (BPI), a potent element in host defense against gram-negative bacteria and lipopolysaccharide. *Immunobiology* **187**, 417-429
- Mori, H., Hara, S., Mizushima, H., Kudo, I., and Inoue, K. (1991) Extracellular PLA<sub>2</sub> detected at inflamed sites in rats does not originate from platelets. *Agents Actions* **32**, 266-269
- Vadas, P., and Pruzanski, W. (1993) Induction of group II phospholipase A<sub>2</sub> expression and pathogenesis of the sepsis syndrome. *Crit. Shock* **39**, 160-167
- Buchler, M., Delier, A., Malferrheiner, P., Kleine, H. O., Wiedeck, H., Uhl, W., Samtner, M., Friess, H., Nevalainen, T., and Beger, H. G. (1989) Serum phospholipase A<sub>2</sub> in intensive care patients with peritonitis, multiple injury, and necrotizing pancreatitis. *Klin. Wochenschr.* **67**, 217-21
- Murakami, M., Kudo, I., Suwa, Y., and Inoue, K. (1992) Release of 14-kDa group-II phospholipase A<sub>2</sub> from activated mast cells and its possible involvement in the regulation of the degranulation process. *Eur. J. Biochem.* **209**, 257-265
- Asaoka, Y., Yoshida, K., Sasaki, Y., Nishizuka, Y., Murakami, M., Kudo, I., and Inoue, K. (1993) Possible role of mammalian secretory group II phospholipase A<sub>2</sub> in T-lymphocyte activation: implication in propagation of inflammatory reaction. *Proc. Natl. Acad. Sci. USA* **90**, 716-719

## REVIEWS

45. Murakami, M., Kudo, I., and Inoue, K. (1993) Molecular nature of phospholipase A<sub>2</sub> involved in prostaglandin I<sub>2</sub> synthesis in human umbilical vein endothelial cells. Possible participation of cytosolic and extracellular type II phospholipases A<sub>2</sub>. *J. Biol. Chem.* 268, 839-844
46. Barbour, S. E., and Dennis, E. A. (1993) Antisense inhibition of group II phospholipase A<sub>2</sub> expression blocks the production of prostaglandin E<sub>2</sub> by P388D1 cells. *J. Biol. Chem.* 268, 21875-21882
47. Arita, H., and Hanasaki, K. (1993) Physiological aspects of a high affinity binding site for pancreatic-type phospholipase A<sub>2</sub>. *J. Lipid Mediat.* 6, 217-222
48. Lin, L. L., Wartmann, M., Lin, A. Y., Knopf, J. L., Seth, A., and Davis, R. J. (1993) cPLA<sub>2</sub> is phosphorylated and activated by MAP kinase. *Cell* 72, 269-278
49. Qiu, Z.-H., de Carvalho, M. S., and Leslie, C. C. (1993) Regulation of phospholipase A<sub>2</sub> activation by phosphorylation in mouse peritoneal macrophages. *J. Biol. Chem.* 268, 24506-24513
50. Kramer, R. M., Roberts, E. F., Manetta, J. V., Sportsman, J. R., and Jakubowski, J. A. (1993) Ca<sup>2+</sup>-sensitive cytosolic phospholipase A<sub>2</sub> (cPLA<sub>2</sub>) in human platelets. *J. Lipid Mediat.* 6, 209-216
51. Lehman, J. J., Brown, K. A., Ramanadham, S., Turk, J., and Gross, R. W. (1993) Arachidonic acid release from aortic smooth muscle cells induced by [Arg8]vasopressin is largely mediated by calcium-independent phospholipase A<sub>2</sub>. *J. Biol. Chem.* 268, 20713-6
52. Jain, M. K., Rogers, J., Jahagirdar, D. V., Mzarecek, J. F., and Ramirez, F. (1986) Kinetics of interfacial catalysis by phospholipase A<sub>2</sub> in intravesicle scooting mode, and heterofusion of anionic and zwitterionic vesicles. *Biochim. Biophys. Acta* 860, 435-447
53. Jain, M. K., Ranadive, G., Yu, B.-Z., and Verheij, H. M. (1991) Interfacial catalysis by phospholipase A<sub>2</sub>: monomeric enzyme is fully catalytically active at the bilayer interface. *Biochemistry* 30, 7330-7340
54. Bayburt, T., Yu, B. Z., Lin, H. K., Browning, J., Jain, M. K., and Gelb, M. H. (1993) Human nonpancreatic secreted phospholipase A<sub>2</sub>: interfacial parameters, substrate specificities, and competitive inhibitors. *Biochemistry* 32, 573-582
55. Hanel, A. M., Schuttel, S., and Gelb, M. H. (1993) Processive interfacial catalysis by mammalian 85-kilodalton phospholipase A<sub>2</sub> enzymes on product-containing vesicles: application to the determination of substrate preferences. *Biochemistry* 32, 5949-5958
56. Jain, M. K., Egmond, M. R., Verheij, H. M., Apitz-Castro, R., Dijkman, R., and de Haas, G. H. (1982) Interaction of phospholipase A<sub>2</sub> and phospholipid bilayers. *Biochim. Biophys. Acta* 688, 341-348
57. Jain, M. K., and Berg, O. (1989) The kinetics of interfacial catalysis by phospholipase A<sub>2</sub> and regulation of interfacial activation: hopping versus scooting. *Biochim. Biophys. Acta* 1002, 127-156
58. Roberts, M. F., Deems, R. A., and Dennis, E. A. (1977) Dual role of interfacial phospholipid in phospholipase A<sub>2</sub> catalysis. *Proc. Natl. Acad. Sci. USA* 74, 1950-1954
59. Volwerk, J. J., Pieterse, W. A., and de Haas, G. H. (1974) Histidine at the active site of phospholipase A<sub>2</sub>. *Biochemistry* 13, 1446-1454
60. Ramirez, F., and Jain, M. K. (1991) Phospholipase A<sub>2</sub> at the bilayer interface. *Protein* 9, 229-239
61. Apitz-Castro, R., Jain, M. K., and de Haas, G. H. (1982) Origin of the latency phase during the action of phospholipase A<sub>2</sub> on unmodified phosphatidylcholine vesicles. *Biochim. Biophys. Acta* 688, 349-356
62. Bell, J. D., and Biltonen, R. L. (1992) Molecular details of the activation of soluble phospholipase A<sub>2</sub> on lipid bilayers: Comparison of computer simulations with experimental results. *J. Biol. Chem.* 267, 11046-11056
63. Jain, M. K., Gelb, M. H., Rogers, J., and Berg, O. G. (1994) The kinetic basis for interfacial catalysis by phospholipase A<sub>2</sub>. *Methods Enzymol.* In press
64. Gelb, M. H., Jain, M. K., and Berg, O. (1992) Interfacial enzymology of phospholipase A<sub>2</sub>. *Bioorg. Med. Chem. Lett.* 2, 1335-1342
65. Jain, M. K., Yu, B.-Z., Rogers, J., Ranadive, G. N., and Berg, O. (1991) Interfacial catalysis by phospholipase A<sub>2</sub>: dissociation constants for calcium, substrate, and competitive inhibitors. *Biochemistry* 30, 7306-7317
66. Jain, M. K., Yuan, W., and Gelb, M. H. (1989) Competitive inhibition of phospholipase A<sub>2</sub> in vesicles. *Biochemistry* 28, 4135-4139
67. Jain, M. K., and Jahagirdar, D. V. (1985) Action of phospholipase A<sub>2</sub> on bilayers. Effect of inhibitors. *Biochim. Biophys. Acta* 814, 319-326
68. Jain, M. K., Ghomashchi, F., Yu, B.-Z., Bayburt, T., Murphy, D., Houck, D., Brownell, J., Reid, J. C., Solowiej, J. E., Wong, S.-M., Moczek, U., Jarrell, R., Sasser, M., and Gelb, M. H. (1992) Fatty acid amides: scooting mode-based discovery of tight-binding competitive inhibitors of secreted phospholipases A<sub>2</sub>. *J. Med. Chem.* 35, 3584-3586
69. Jain, M. K., Yu, B.-Z., Gelb, M. H., and Berg, O. G. (1992) Assay of phospholipase A<sub>2</sub> and their inhibitors by kinetic analysis in the scooting mode. *Med. Inflamm.* 1, 85-100
70. Ransac, S., Riviere, C., Soulie, J. M., Gancet, C., Verger, R., and de Haas, G. H. (1990) Competitive inhibition of lipolytic enzymes. *Biochim. Biophys. Acta* 1043, 57-66
71. Yu, L., and Dennis, E. A. (1991) Thio-based phospholipase assay. *Methods Enzymol.* 197, 65-75
72. Yu, L., and Dennis, E. A. (1991) Critical role of a hydrogen bond in the interaction of phospholipase A<sub>2</sub> with transition-state and substrate analogues. *Proc. Natl. Acad. Sci. USA* 88, 9325-9329
73. Street, I. P., Lin, H. K., Laliberte, F., Ghomashchi, F., Wang, Z., Perrier, H., Tremblay, N. M., Huang, Z., Weech, P. K., and Gelb, M. H. (1993) Slow- and tight-binding inhibitors of the 85-kDa human phospholipase A<sub>2</sub>. *Biochemistry* 32, 5935-5940
74. Jain, M. K., Rogers, J., Hendrickson, H. S., and Berg, O. G. (1993) The chemical step is not rate-limiting during the hydrolysis by phospholipase A<sub>2</sub> of mixed micelles of phospholipid and detergent. *Biochemistry* 32, 8360-8367
75. Jain, M. K., Rogers, J., Berg, O., and Gelb, M. H. (1991) Interfacial catalysis by phospholipase A<sub>2</sub>: activation by substrate replenishment. *Biochemistry* 30, 7340-7348
76. Dijkman, R., Lugtigheid, R. B., Dekker, N., Van, D. B. L., Egmond, M. R., and Verheij, H. M. (1993) Competitive inhibition of lipolytic enzymes. IX. A comparative study on the inhibition of pancreatic phospholipases A<sub>2</sub> from different sources by (R)-2-acylamino phospholipid analogues. *Biochim. Biophys. Acta* 1167, 281-288
77. Yuan, W., and Gelb, M. H. (1988) Phosphonate-containing phospholipid analogues as tight-binding inhibitors of PLA<sub>2</sub>. *J. Am. Chem. Soc.* 110, 2665-2666
78. Jain, M. K., Yuan, W., and Gelb, M. H. (1989) Competitive inhibition of phospholipase A<sub>2</sub> in vesicles. *Biochemistry* 28, 4135-4139
79. Jain, M. K., Tao, W., Rogers, J., Arenson, C., Eibl, H., and Yu, B.-Z. (1991) Active-site specific competitive inhibitors of PLA<sub>2</sub>: novel transition state analogs. *Biochemistry* 30, 10256-10268
80. Plesniak, L. A., Boegeman, S. C., Segelke, B. W., and Dennis, E. A. (1993) Interaction of phospholipase A<sub>2</sub> with thioether amide containing phospholipid analogues. *Biochemistry* 32, 5009-5016
81. Trimble, L. A., Street, I. P., Perrier, H., Tremblay, N. M., Weech, P. K., and Bernstein, M. A. (1993) NMR structural studies of the tight complex between a trifluoromethyl ketone inhibitor and the 85-kDa human phospholipase A<sub>2</sub>. *Biochemistry* 32, 12560-12565
82. Morgan, B. P., Scholtz, J. M., Ballinger, M. D., Zipkin, I. D., and Bartlett, P. A. (1991) Differential binding energy: a detailed evaluation of the influence of hydrogen-bonding and hydrophobic groups on the inhibition of thermolysin by phosphorus-containing inhibitors. *J. Am. Chem. Soc.* 113, 297-307
83. Lin, H.-K., and Gelb, M. H. (1993) Competitive inhibition of interfacial catalysis by phospholipase A<sub>2</sub>: differential interaction of inhibitors with the vesicle interface as a controlling factor of inhibitor potency. *J. Am. Chem. Soc.* 115, 3932-3942
84. Jain, M. K., Yu, B. Z., and Berg, O. G. (1993) Relationship of interfacial equilibria to interfacial activation of phospholipase A<sub>2</sub>. *Biochemistry* 32, 11319-11329
85. Yuan, W., Quinn, D. M., Sigler, P. B., and Gelb, M. H. (1990) Kinetic and inhibition studies of phospholipase A<sub>2</sub> with short-chain substrates and inhibitors. *Biochemistry* 29, 6082-6094
86. Bukowski, T., and Teller, D. C. (1986) Self-association and active enzyme forms of Naja naja naja and Crotalus atrox phospholipases A<sub>2</sub> studied by analytical ultracentrifugation. *Biochemistry* 25, 8024-8033
87. Pluckhahn, A., and Dennis, E. A. (1985) Activation, aggregation, and product inhibition of cobra venom phospholipase A<sub>2</sub> and comparison with other phospholipases. *J. Biol. Chem.* 260, 11099-11106
88. van Eijk, J. H., Verheij, H. M., Dijkman, R., and de Haas, G. H. (1983) Interaction of phospholipase A<sub>2</sub> from Naja melanoleuca snake venom with monomeric substrate analogs: Activation of the enzyme by protein-protein or lipid-protein interactions? *Eur. J. Biochem.* 132, 183-188
89. van Oort, M. G., Dijkman, R., Hille, J. D., and de Haas, G. H. (1985) Kinetic behavior of porcine pancreatic phospholipase A<sub>2</sub> on zwitterionic and negatively charged double-chain substrates. *Biochemistry* 24, 7993-7999
90. Rogers, J., Yu, B. Z., and Jain, M. K. (1992) Basis for the anomalous effect of competitive inhibitors on the kinetics of hydrolysis of short-chain phosphatidylcholines by phospholipase A<sub>2</sub>. *Biochemistry* 31, 6056-6062
91. Bartoli, F., Lin, H.-K., Ghomashchi, F., Gelb, M. H., Jain, M. K., and Apitz-Castro, R. (1994) Tight-binding inhibitors of 85 kDa phospholipase A<sub>2</sub> but not 14 kDa phospholipase A<sub>2</sub> inhibit the release of arachidonate in thrombin-stimulated human platelets. *J. Biol. Chem.* 269, 15625-15630
92. Riendeau, D., Guay, J., Weech, P. K., Laliberte, F., Yergey, J., Li, C., Desmarais, S., Perrier, H., Liu, S., Nicoll-Griffith, D., and Street, I. P. (1994) Arachidonyl trifluoromethyl ketone, a potent inhibitor of 85-kDa phospholipase A<sub>2</sub>, blocks production of arachidonate and 12-hydroxyeicosatetraenoic acid by calcium ionophore-challenged platelets. *J. Biol. Chem.* 269, 1-6
93. Gronich, J., Konieczkowski, M., Gelb, M. H., Nemenoff, R. A., and Sedor, J. R. (1994) Interleukin 1α causes rapid activation of cytosolic phospholipase A<sub>2</sub> by phosphorylation in rat mesangial cells. *J. Clin. Invest.* 93, 1224-1233
94. Fisher, A. B., Dodia, C., Chander, A., and Jain, M. (1992) A competitive inhibitor of phospholipase A<sub>2</sub> decreases surfactant phosphatidylcholine degradation by the rat lung. *Biochem. J.* 288, 407-411

## Periodicity of leucine and tandem repetition of a 24-amino acid segment in the primary structure of leucine-rich $\alpha_2$ -glycoprotein of human serum

(protein structure/amphiphilic structure/gene duplication/membrane protein/galactosamine and glucosamine oligosaccharides)

NOBUHIRO TAKAHASHI, YOKO TAKAHASHI, AND FRANK W. PUTNAM\*

Department of Biology, Indiana University, Bloomington, IN 47405

Contributed by Frank W. Putnam, November 15, 1984

**ABSTRACT** The complete primary structure of the 3.1S leucine-rich  $\alpha_2$ -glycoprotein (LRG) present in human plasma has been determined. This protein ( $M_r \approx 45,000$ ) consists of a single polypeptide chain with one galactosamine and four glucosamine oligosaccharides attached. The polypeptide has two intrachain disulfide bonds and contains 312 amino acid residues of which 66 are leucine. The amino acid sequence can be exactly divided into 13 segments of 24 residues each, eight of which exhibit a periodic pattern in the occurrence of leucine, proline, and asparagine. The consensus sequence for the repeating tetracosapeptide unit is Pro-Xaa-Xaa-Leu-Leu-Xaa-Xaa-Xaa-Xaa-Leu-Xaa-Xaa-Leu-Xaa-Leu-Xaa-Xaa-Asn-Xaa-Leu-Xaa-Xaa-Leu. This periodicity suggests that the unique structure of LRG arose from a series of unequal crossovers of a precursor oligonucleotide sequence that encoded a building block rich in leucine. Overall, the amino acid sequence of LRG is not significantly homologous to the continuous sequence of any protein in the current data base. However, the consensus tetracosapeptide sequence shows strong homology to segments of many mitochondrial proteins, viral envelope proteins, and oncogene proteins that have a high leucine content and transmembrane domains. Tandem repetition of similar segments also occurs in apolipoproteins that have amphipathic helical potential. Prediction of the secondary structure by the Chou-Fasman rules and calculation of the hydrophilic/hydrophobic profile by several methods confirm the tandem repetition of largely hydrophobic structural units; these begin with a  $\beta$ -turn that leads into an organized structure with  $\alpha$ -helical or  $\beta$ -sheet potential. These structural characteristics and the homology to mitochondrial proteins and apolipoproteins suggest that LRG is a membrane-derived or membrane-associated protein containing a series of domains capable of bipolar surface orientation.

Although complete amino acid sequences have been reported for about half of the nearly 100 proteins that have been isolated from human plasma (1), little is known about the structure of a series of well-characterized  $\alpha$ - and  $\beta$ -glycoproteins of uncertain or unknown function (2). We have begun study of some of these as part of a program for investigation of the structure and function of human plasma proteins (3, 4). We report here the complete amino acid sequence of a trace component of unknown function that was isolated from human serum in 1977 by Haupt and Baudner (5). They called it 3.1S leucine-rich  $\alpha_2$ -glycoprotein (LRG) because of its sedimentation coefficient, its unusually high content of leucine, and its electrophoretic mobility. The protein was reported to consist of a single polypeptide chain ( $M_r \approx 50,000$ ) and to contain about 23% carbohydrate by weight (5). To our

knowledge, no other structural data have since been published.

LRG is a trace protein of human plasma. The average content of LRG in adult serum is 2.1 mg/100 ml (5). Although neither the site of synthesis nor the function are known and no relationship to disease has been reported, LRG is of interest because of its unusually high content of leucine, which is about 17% by weight. Thus, about every fifth residue is leucine. Whether the leucine is randomly distributed or occurs in a regular structural pattern could only be determined by amino acid sequencing, which we undertook and completed. We found that the sequence exhibits a novel periodic pattern unlike that previously reported for soluble globular proteins except apolipoproteins (6-8). The single polypeptide chain is composed of 312 amino acid residues, and the sequence can be exactly divided into 13 blocks of 24 residues; eight of these tetracosapeptide sequences exhibit a periodic pattern in the distribution of leucine and certain other residues. The periodic pattern is highly significant statistically; the segment comparison score in SD units for the top 100 scores is 18.06. This suggests that LRG arose by a series of unequal crossovers of an oligonucleotide sequence that encoded a prototype leucine-rich building block of about 24 amino acid residues.

The periodicity in the primary structure of LRG is due to a recurring pattern of leucine residues that produces a segment similar in hydrophobicity and length to that found in the transmembrane segments of amphiphilic (amphipathic) peptides and proteins (9, 10). Thus, we suggest that LRG may be a membrane-associated or membrane-derived protein. This hypothesis is supported by predictions of the secondary structure and the hydrophilic/hydrophobic character (hydropathy profile) of the polypeptide chain.

### MATERIALS AND METHODS

**Materials.** Purified LRG prepared from human serum by the method of Haupt and Baudner (5) and antiserum specific for LRG were obtained from Behringwerke AG (Marburg/Lahn, F.R.G.). The protein migrated as a single band in Na-DodSO<sub>4</sub>/polyacrylamide gel electrophoresis both in the presence and absence of 2-mercaptoethanol.

**Methods.** Methods used for the determination of the primary structure of proteins in our laboratory have been described (3, 4, 11). The purified LRG was reduced and carboxymethylated before sequence analysis. The carboxymethylated protein (5-15 mg) was subjected to separate digestions with L-1-tosylamido-2-phenylethyl chloromethyl ketone-treated trypsin, *Staphylococcus aureus* V8 protease, and endoproteinase Lys-C. The intact protein was also

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: LRG, leucine-rich  $\alpha_2$ -glycoprotein.  
\*To whom reprint requests should be addressed.



cleaved with dilute acid at 108°C for 4 hr (12). Each enzymatic digest was separated by a combination of gel filtration and high-performance liquid chromatography (HPLC) on an Ultrasphere ODS column (Altex, Berkeley, CA) or a Synchropak RP-P column (SynChrom, Linden, IN) with a programmed gradient of 1-propanol containing 0.1% trifluoroacetic acid or 0.1% heptafluorobutyric acid (4). The digest with dilute acid was separated by an automated two-step HPLC system; this consists of a combination of anion-exchange chromatography with a Spherogel-TSK IEX-540 DEAE column (Altex) as the first step and chromatography with a reversed-phase Ultrasphere ODS column as the second. The purified peptides were analyzed with the amino acid analyzer, and their sequences were determined by automatic Edman degradation (3, 4, 11). Hexosamine analysis was also done with the amino acid analyzer, after acid hydrolysis (3, 11).

**Computer Analysis of Sequence Data.** The sequence data base of the *Atlas of Protein Sequence and Structure* updated to August 1984 and the programs SEARCH, ALIGN, RELATE, and PRPLOT were provided by the National Biomedical Research Foundation.<sup>†</sup> The programs SEARCH, RELATE, and ALIGN (13) were used either with the unitary matrix or with the mutation data matrix; all gave a score for statistical significance in standard deviations (SD) of the real score above a score of 100 random runs. We used PRPLOT to plot the hydrophilic/hydrophobic profile of the molecule by utilizing different scales including those of Kyte and Doolittle (14), Wolfenden *et al.* (15), Argos *et al.* (16), and Eisenberg (10). PRPLOT was also used to estimate the secondary structure predicted by the procedure of Chou and Fasman (17) based on their values for the tendency of individual amino acids to appear in  $\alpha$ -helix,  $\beta$ -sheet, and  $\beta$ -turn structures.

## RESULTS AND DISCUSSION

**Amino Acid Composition.** The amino acid composition calculated from the sequence analysis (Table 1) corresponds well with that obtained by amino acid analysis of the protein (5). The very high content of leucine (66 of 312 residues) is an unusual characteristic of this protein. By amino acid analysis the leucine content of the glycoprotein is 16.62% by weight (5). The value for the polypeptide portion calculated from the complete sequence is 21.75% by weight or 21.12 mole %. A search of the protein sequence data base<sup>†</sup> showed that the leucine content of LRG is exceeded only by those of a vespid venom peptide, a few hypothetical bacteriophage and mitochondrial proteins corresponding to gene sequences, and the hormone secretin. Most other sequences listed that had a high leucine content (15–20 mole %) also were for hypothetical mitochondrial proteins or were for known mitochondrial cytochromes and enzymes, viral coat proteins, or biologically active peptides. About one-third of the amino acid residues of LRG are strongly hydrophobic. Other notable features are the low content of tyrosine (3 residues), isoleucine (4 residues), and cysteine (4 residues). The latter are linked in two disulfide bonds; one is close to the amino terminus and is between cysteine-8 and cysteine-21, and the other is near the carboxyl terminus and is between cysteine-268 and cysteine-294.

**Polypeptide Structure and Molecular Weight.** Human LRG consists of a single polypeptide chain containing 312 amino acid residues, one GalN oligosaccharide, and four GlcN oli-

Table 1. Amino acid composition of human LRG based on the complete amino acid sequence determination

Amino acid	No. of residues	Amino acid	No. of residues
Aspartic acid	19	Valine	14
Asparagine	17	Methionine	3
Threonine	12	Isoleucine	4
Serine	21	Leucine	66
Glutamic acid	14	Tyrosine	3
Glutamine	20	Phenylalanine	10
Proline	22	Lysine	12
Glycine	23	Histidine	7
Alanine	21	Arginine	15
Half-cystine	4	Tryptophan	5

$M_r$  of unmodified polypeptide chain is 34,346; number of residues is 312. Thr-2 is linked to a GalN oligosaccharide. Asn-44, Asn-151, Asn-234, and Asn-290 are linked to GlcN oligosaccharides.

gosaccharides (Fig. 1). The  $M_r$  calculated for the unglycosylated polypeptide chain is 34,346. An  $M_r$  of 35,970 was obtained from electrophoresis in NaDodSO<sub>4</sub>/polyacrylamide gels when the results estimated in gels ranging from 5% to 20% gel concentration were extrapolated by use of a double-reciprocal plot. This value, which presumably reflects only the polypeptide chain, is much less than the  $M_r$  of 49,600  $\pm$  4,000 determined by equilibrium sedimentation (5). However, the carbohydrate content of LRG (about 23% by weight) contributes significantly to the  $M_r$  and also makes it difficult to determine the  $M_r$  accurately by physicochemical methods. We estimate that the  $M_r$  of LRG is approximately 45,000, based on the value of 34,346 for the polypeptide portion and an estimate of 600 for the GalN oligosaccharide and of 2500 for each of the four GlcN oligosaccharides.

**Number and Location of Oligosaccharides.** Fig. 1 gives the complete amino acid sequence of LRG and shows the sites of attachment of the five oligosaccharides. The GalN is O-linked to the second residue (Thr-2). Both in its proximity to the amino terminus and to a nearby proline, this linkage is similar to that of the GalN attached to the amino-terminal residue (threonine) in human hemopexin (11, 18). However, no obligatory signal sequence for the attachment of GalN is known, whereas GlcN is always N-linked to asparagine in the tripeptide sequence Asn-Xaa-(Thr/Ser). In LRG there are four GlcN oligosaccharides; these are linked to Asn-44, Asn-151, Asn-234, and Asn-290. However, the asparagine is not glycosylated in the putative acceptor sequence Asn-Leu-Ser at positions 271–273. Nothing has been published about the carbohydrate structure of the oligosaccharides of LRG; however, a complex dibranched GlcN oligosaccharide is the most common carbohydrate in human plasma glycoproteins (19).

**Internal Homology.** Although many plasma proteins exhibit strong evidence of internal duplication in amino acid sequence (20, 21), at first appearance LRG does not. Despite the frequency of leucine residues, there are only five instances of a Leu-Leu sequence and only one case of the triplet sequence Leu-Leu-Leu, compared to 53 instances of single leucine residues. This lack of a statistical distribution puzzled us. However, first by visual inspection and then by computer analysis with the program RELATE, we found that the amino acid sequence of LRG fits a 13-cycle periodic pattern dominated by a 24-residue peptide (tetracosapeptide) which has the consensus sequence Pro-Xaa-Xaa-Leu-Leu-Xaa-Xaa-Xaa-Xaa-Xaa-Leu-Xaa-Xaa-Leu-Xaa-Leu-Xaa-Xaa-Asn-Xaa-Leu-Xaa-Xaa-Leu. All of the amino acid residues identified in this sequence occur at least six times in the alignment shown in Fig. 2. In the computer listing of the 100 segments having the highest scores for intrasequence homology (the top 100 of 39,903 comparisons), the lengths of all

<sup>†</sup>Barker, W. C., Hunt, L. T., Orcutt, B. C., George, D. G., Yeh, L. S., Chen, H. R., Blomquist, M. C., Johnson, G. C. & Dayhoff, M. O. (Aug. 1, 1984) *Atlas of Protein Sequence and Structure*, Protein Sequence Database, (Natl. Biomed. Res. Found., Washington, D.C.).

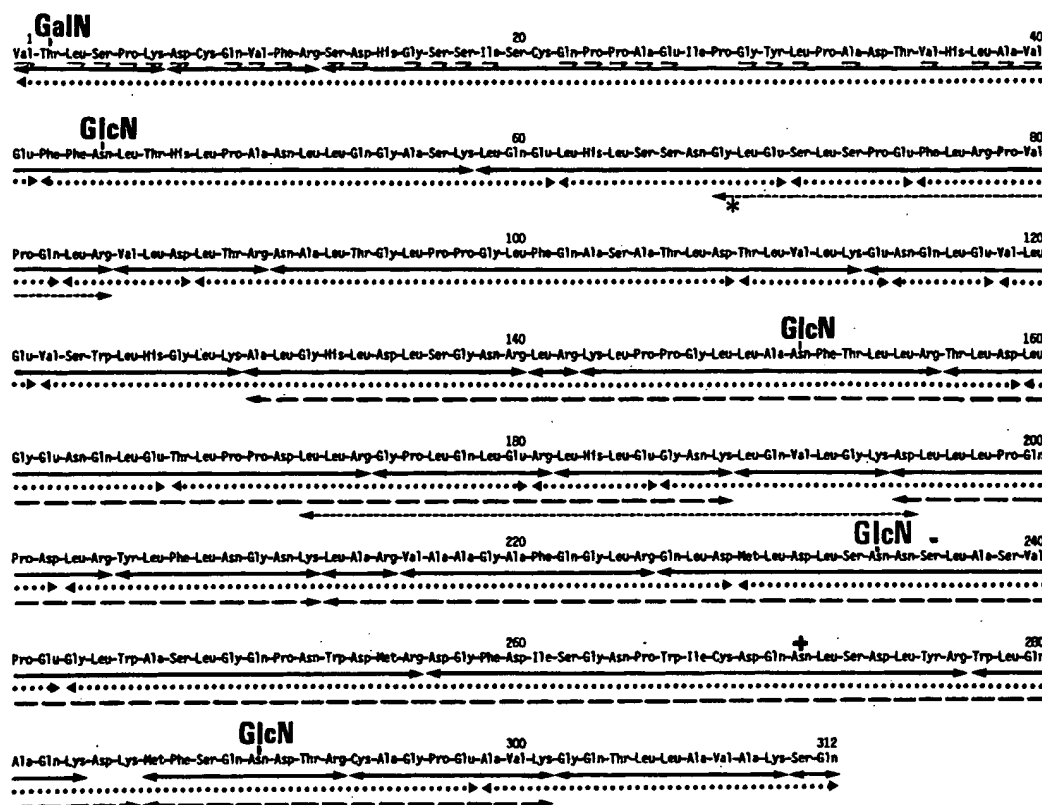


FIG. 1. Summary of the complete amino acid sequence of human LRG. The sequence is shown along with all peptides necessary for the proof of sequence. The peptides obtained from different digestions are as follows: —, tryptic peptides; - - - - - , *S. aureus* V8 peptides; . . . . . , endoproteinase Lys-C peptides; - · - · - · , dilute acid peptides. The asterisk indicates a dilute acid peptide of a tryptic peptide. GalN and GlcN indicate the attachment sites for the galactosamine oligosaccharide and the glucosamine oligosaccharides, respectively. A potential glycosylation site (Asn-271) is shown by a plus sign: No evidence for polymorphism in the sequence of LRG was found except for a possible ambiguity at position 215, where tyrosine may replace arginine in some molecules.

segments listed were divisible by 24, and all the segments fit exactly into the periodic pattern of Fig. 2. In this pattern, the complete sequence of 312 residues is exactly divided into an array of 13 tetracosapeptide segments with no gaps inserted. The segment comparison score for the tetracosapeptide segments is 18.06 SD units. A score of 3.0 SD units is considered significant (13).

The unique feature of the 13-cycle alignment is that it is

based mainly on a periodic distribution of a single amino acid rather than on a recurrent sequence of several different amino acids. More than 50 of the 66 leucine residues fit the pattern, including almost all the leucines in the 200 successive residues (nos. 38–237) in which the most homologous tetracosapeptide segments are located. Four more leucines are aligned in column 8 of Fig. 2. Since leucine is one of the most hydrophobic amino acids and is usually present in an or-

1	1	V	T	L	S	P	K	D	C	Q	V	F	R	S	D	H	G	S	S	I	S	C	Q	P	P	24
2	25	A	E	I	P	G	Y	L	P	A	D	T	V	H	A	V	E	F	F	N	*	T	H	E	S	48
3	49	A	N	E	Q	G	A	S	K	Q	E	L	H	S	S	S	S	G	A	E	T	G	E	S	72	
4	73	S	P	E	F	R	P	V	P	Q	R	V	L	D	E	T	R	N	A	E	T	G	E	S	96	
5	97	P	G	E	F	Q	A	S	A	T	D	T	V	K	E	N	Q	E	V	E	V	E	V	E	120	
6	121	E	V	S	W	H	G	L	K	A	G	H	D	E	S	G	M	R	E	R	K	E	S	144		
7	145	P	G	E	F	Q	A	S	A	T	D	T	V	K	E	N	Q	E	V	E	V	E	V	168		
8	169	P	D	L	R	G	P	L	Q	E	R	H	E	G	N	K	Q	V	E	S	192					
9	193	G	K	D	L	P	Q	P	D	R	Y	F	N	G	K	L	A	R	V	216						
10	217	A	A	G	A	F	Q	G	L	R	Q	D	M	D	L	S	N	S	I	A	S	V	240			
11	241	E	G	W	A	S	L	G	Q	P	N	W	D	M	R	D	G	F	D	I	S	G	N	264		
12	265	W	I	C	D	N	L	S	D	E	Y	R	W	L	Q	A	Q	K	D	K	M	F	S	288		
13	289	Q	N	D	T	R	C	A	G	P	E	A	V	K	G	Q	T	L	L	A	V	A	K	S	312	
Consensus		P	P	G	L	L	Q	G	L	P	Q	L	R	X	L	D	L	S	G	N	X	L	E	S	L	

FIG. 2. Periodicity of leucine residues and tandem repetition of a 24-amino acid segment in the primary structure of LRG. The sequence of Fig. 1 with position numbers is shown in the one-letter code for amino acids (13) and is aligned exactly into 13 segments (numbered in bold type). Amino acids that occur at least six times in a vertical column are shown in bold type and are enclosed in shaded boxes. The consensus sequence is given at the bottom, with the most frequent residues in bold type. Others that occur three or four times are also given in the consensus sequence, and X is used where there is no clear preference. The cysteine (C) residues are circled and the disulfide bonds are represented schematically. An asterisk denotes each asparagine (N) to which a GlcN oligosaccharide is linked, and a solid circle at threonine-2 indicates that a GalN oligosaccharide is attached.

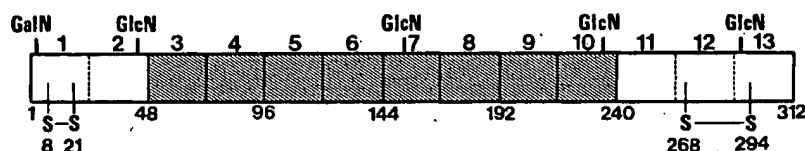


FIG. 3. Schematic representation of the tandem repetition of 24-amino acid segments in LRG. Each numbered block represents a segment (see Fig. 2). Shaded blocks represent homologous segments that exhibit a periodic pattern of leucine and other amino acids. The numbers below the diagram give the sequence positions. The locations of the GalN and GlcN oligosaccharides and of the two disulfide bonds are shown.

dered structure ( $\alpha$ -helix or  $\beta$ -pleated sheet), the periodicity of leucine residues probably reflects a unique structural characteristic of LRG that has been attained by a series of unequal crossovers of an oligonucleotide sequence that coded for a prototype building block that was rich in leucine. Although the entire polypeptide sequence is exactly divisible into 13 tetracosapeptide segments, only the central 8 segments exhibit strong inter-segment homology. These comprise a central core structure, illustrated schematically in Fig. 3. The amino-terminal segment has a GalN oligosaccharide and an intra-segment disulfide bond. The periodic pattern of homology ends after segment 10. The last two segments are joined by a disulfide bond, and the last has a GlcN oligosaccharide.

**Sequence Homology to Other Proteins.** To examine the possible occurrence in other proteins of sequences homologous to the periodic tetracosapeptide of LRG, we used as a probe a consensus sequence that represents the amino acid of highest frequency in each vertical column of Fig. 2. The sequence used was Pro-Pro-Gly-Leu-Leu-Gln-Gly-Leu-Pro-Gln-Leu-Arg-Xaa-Leu-Asp-Leu-Ser-Gly-Asn-Xaa-Leu-Glu-Ser-Leu, in which Xaa indicates that no amino acid occurred with sufficient frequency.

By use of the program SEARCH and both the unitary matrix and the mutation data matrix, we found that LRG has a unique amino acid sequence that does not exhibit significant homology over a series of contiguous segments to any protein in the current sequence data base.<sup>†</sup> However, most of the repeating tetracosapeptide segments of LRG do exhibit striking homology of their basic pattern with segments of a number of polypeptides, most of which have a high leucine content. These polypeptides include many viral envelope proteins, oncogene products, and a number of known mitochondrial proteins, such as various cytochromes as well as hypothetical mitochondrial proteins for which the sequences have been deduced by cloning techniques. For example, when the unitary matrix was used, the highest score for the prototype tetracosapeptide sequence was given by a 24-residue segment of hypothetical protein 4 of the human mitochondrion, which has a leucine content of 21 mole % (22). Other segments of protein 4 and other mitochondrial proteins scored high, but the matching segments were not contiguous in the sequences of the proteins. When the mutation data matrix was used, the highest score was for the hemagglutinin precursor of various strains of human and duck influenza virus (residues 441–464). This segment corresponds to residues 78–101 in the second (HA<sub>2</sub>) chain of the processed virus protein and exists as a long  $\alpha$ -helix with a 36-Å hydrophobic exterior surface (23).

We do not suggest that the homology identified signifies a genetic or evolutionary relationship. Rather, it probably reflects the fact that a high leucine content may result in a periodicity in sequence similar to that of LRG. In an organized structure such as the  $\alpha$ -helix or  $\beta$ -pleated sheet (for both of which leucine has a high potential), the hydrophobic residues tend to cluster in a similar orientation (9, 10). Such hydrophobic patches must either be in the interior of protein molecules or, if on the exterior, will tend to interact with hydrophobic membrane surfaces.

The high solubility of LRG and its resistance to heat coagulation are probably due to the combination of a high carbohydrate content, which greatly increases the hydrophilicity of the protein surface, and to the clustering of the hydrophobic leucine residues, which are probably internalized in the protein.

**Tandem Repetitions of Protein Sequences.** Although fibrous proteins such as collagen have frequent short (tripeptide) repeating units, tandem repetitions of the length and type found in LRG are rare. However, some apolipoproteins contain a series of 11-amino acid or 22-amino acid segments similar to although not closely homologous with the repeating sequence in LRG (6–8). Fitch (6) identified a 13-cycle repeat of an 11-amino acid (or 22-amino acid) sequence in human apolipoprotein A-I and proposed that it resulted from a series of intragenic unequal crossovers. Barker and Dayhoff (7) identified a similar periodicity in human apolipoproteins A-II, C-II, and C-III and suggested that the four lipoproteins arose from a single ancestral gene. Recently Boguski *et al.* (8) reported that rat apolipoprotein A-IV contains 13 tandem repetitions of a 22-amino acid segment with amphipathic helical potential. The consensus sequence of the apolipoproteins is similar to that of LRG in length, in the number of repetitions, and in the frequency of leucine and other hydrophobic residues. However, the sequence homology of LRG and apolipoproteins is not pronounced; hence, they probably have a functional rather than a genetic relationship. The apolipoprotein repeating sequence can be drawn as an  $\alpha$ -helical wheel with amphiphilic character; that is, the nonpolar residues are segregated on one section of the circumference, and hydrophilic and charged residues are aligned on the rest. The Chou-Fasman rules (17) lead to the prediction of an alternating secondary structure, which is in accord with a plot of the hydropathic index. The result is a helical amphiphilic structure to which the function of lipid binding is attributed.

**Hydrophilic/Hydrophobic Profile.** When the computer program PRPLOT was used with different scales (10, 14–16) to evaluate the hydrophilic/hydrophobic profile of LRG, the results were in general agreement. Computation of the hydropathic index of Kyte and Doolittle (14) (Fig. 4) shows that the eight conserved segments of the central core are largely hydrophobic, whereas the amino-terminal segment and the last two segments are mainly hydrophilic. Each of the central eight segments has a predominantly hydrophobic pattern with a short hydrophilic section in the middle or near the carboxyl terminus. The patterns for the eight tandem segments are similar but not identical. The carboxyl-terminal segment appears to end with a hydrophobic segment; this could be buried inside the molecule, or it could be a transmembrane segment. In view of the high leucine content, the hydrophobicity of LRG is not unexpected. However, using the method of Eisenberg (10), we were unable to identify a series of potential membrane segments of 21–24 residues in length, such as those he described in some membrane proteins. Also, the periodic leucine residues did not all cluster together in an  $\alpha$ -helical wheel as the hydrophobic residues tend to do in apolipoproteins and amphiphilic peptides (8–10).

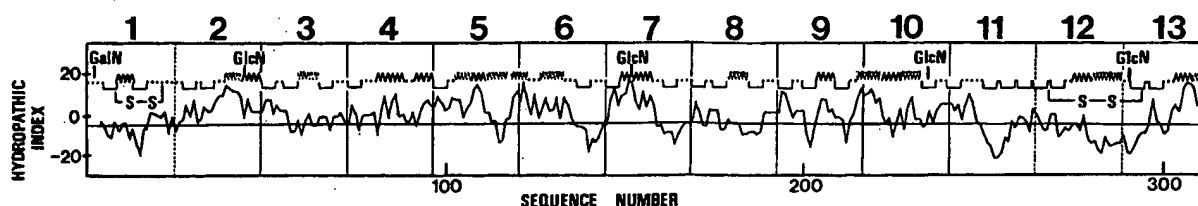


FIG. 4. Hydropathy profile (14) and secondary structure predictions (17) for LRG calculated by use of the PRPLOT program (see text footnote †). In the secondary structure prediction, the residues are represented by dots and are shown in  $\alpha$ -helical (wavy),  $\beta$ -sheet (zigzag), and  $\beta$ -turn (turn) conformational states. Locations are given for the oligosaccharides (GalN and GlcN) and for the disulfide bonds. The vertical lines divide the structure into the 13 numbered tandem segments for which the amino acid sequences are given in Fig. 2.

**Secondary Structure.** The secondary structure predicted by the Chou-Fasman rules (17) shows some repetition of structural elements in the eight tandem conserved segments of the central core of the molecule (shaded section of Fig. 3). In fact, the predicted structure is nearly identical for segments 3 and 8, which have 11 residues in common. Although the predicted secondary structure of the core segments is not identical, the general pattern is similar. Most of the conserved segments begin with a  $\beta$ -turn, followed by an organized structure (either  $\alpha$ -helical or  $\beta$ -sheet) in the middle of the segment, and end with a  $\beta$ -turn. The two segments at the amino terminus and the two at the carboxyl terminus are each the loci of a disulfide bond and of carbohydrate, both of which greatly affect the local structure in a way that is hard to predict by the Chou-Fasman rules. These structural characteristics and the homology to mitochondrial proteins and apolipoproteins suggest that LRG is a membrane-derived or membrane-associated protein containing a series of domains capable of bipolar surface orientation.

We thank P. H. Davidson, S. A. Dorwin, J. Madison, and K. Huss for excellent technical assistance. This work was supported by Grant AM19221 from the National Institutes of Health.

- Putnam, F. W., ed. (1984) *The Plasma Proteins* (Academic, Orlando, FL), 2nd Ed., Vol. 4.
- Schwick, H. G. & Haupt, H. (1984) in *The Plasma Proteins*, ed. Putnam, F. W. (Academic, Orlando, FL), 2nd Ed., Vol. 4, pp. 167-220.
- Lozier, J., Takahashi, N. & Putnam, F. W. (1984) *Proc. Natl. Acad. Sci. USA* 81, 3640-3644.
- Takahashi, N., Takahashi, Y., Ortel, T. L., Lozier, J., Ishioka, N. & Putnam, F. W. (1984) *J. Chromatogr.*, 317, 11-25.
- Haupt, H. & Baudner, S. (1977) *Hoppe-Seyler's Z. Physiol. Chem.* 358, 639-646.
- Fitch, W. M. (1977) *Genetics* 86, 623-644.
- Barker, W. C. & Dayhoff, M. O. (1977) *Comp. Biochem. Physiol. B* 57, 309-315.
- Boguski, M. S., Elshourbagy, N., Taylor, J. M. & Gordon, J. I. (1984) *Proc. Natl. Acad. Sci. USA* 81, 5021-5025.
- Kaiser, E. T. & Kézdy, F. J. (1983) *Proc. Natl. Acad. Sci. USA* 80, 1137-1143.
- Eisenberg, D. (1984) *Annu. Rev. Biochem.* 53, 595-623.
- Takahashi, N., Takahashi, Y. & Putnam, F. W. (1984) *Proc. Natl. Acad. Sci. USA* 81, 2021-2025.
- Inglis, A. S. (1983) *Methods Enzymol.* 91, 324-332.
- Barker, W. C., Ketcham, L. K. & Dayhoff, M. O. (1978) in *Atlas of Protein Sequence and Structure*, Dayhoff, M. O. ed. (Natl. Biomed. Res. Found., Washington, DC), Vol. 5, Suppl. 3, pp. 359-362.
- Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* 157, 105-132.
- Wolfenden, R., Anderson, L., Cullis, P. M. & Southgate, C. C. B. (1981) *Biochemistry* 20, 849-855.
- Argos, P., Rao, J. K. M. & Hargrave, P. A. (1982) *Eur. J. Biochem.* 128, 565-575.
- Chou, P. Y. & Fasman, G. O. (1978) *Annu. Rev. Biochem.* 47, 251-276.
- Takahashi, N., Takahashi, Y. & Putnam, F. W. (1985) *Proc. Natl. Acad. Sci. USA* 82, 73-77.
- Baenziger, J. (1984) in *The Plasma Proteins*, ed. Putnam, F. W. (Academic, Orlando, FL), 2nd Ed., Vol. 4, pp. 271-315.
- Doolittle, R. F. (1984) in *The Plasma Proteins*, ed. Putnam, F. W. (Academic, Orlando, FL), 2nd Ed., Vol. 4, pp. 317-359.
- Ortel, T. L., Takahashi, N. & Putnam, F. W. (1984) *Proc. Natl. Acad. Sci. USA* 81, 4761-4765.
- Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R. & Young, I. G. (1981) *Nature (London)* 290, 457-465.
- Wilson, I. A., Skehel, J. J. & Wiley, D. C. (1981) *Nature (London)* 289, 366-373.

SEQ ID NO: 262

RESULT 8

AAV87784

ID AAV87784 standard; cDNA; 686 BP.

XX

AC AAV87784;

XX

DT 12-FEB-1999 (first entry)

XX

DE EST clone ES35.

XX

KW Expressed sequence tag; secreted protein; haematopoiesis regulator;  
KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;  
KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;  
KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.

XX

OS Homo sapiens.

XX

PN WO9845437-A2.

XX

PD 15-OCT-1998.

XX

PF 10-APR-1998; 98WO-US06956.

XX

PR 10-APR-1997; 97US-0837312.

XX

PA (GEMY ) GENETICS INST INC.

XX

PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

PI Racie LA, Spaulding V, Treacy M;

XX

DR WPI; 1999-070078/06.

XX

PT New polynucleotides encoding human secreted proteins - derived from  
PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,  
PT ovary, pituitary, retina and colon cDNA libraries

XX

PS Claim 1; Page 178; 641pp; English.

XX

CC The present sequence represents an expressed sequence tag (EST), and is  
CC a polynucleotide of the invention. The polynucleotides of the invention  
CC are all secreted EST sequences isolated from a variety of human tissue  
CC sources. The EST sequences and proteins encoded by them are predicted to  
CC have useful biological activities which would make them suitable for  
CC treating, preventing or ameliorating medical conditions in humans and  
CC animals, although no supporting data is given. Suggested activities  
CC include nutritional activity, immune stimulating or suppressing activity,  
CC haematopoiesis regulating activity, tissue growth activity,  
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition  
CC activity. The EST sequences are also stated to be useful for gene  
CC therapy.

XX

SQ Sequence 686 BP; 137 A; 243 C; 177 G; 129 T; 0 other;

Query Match 32.9%; Score 603; DB 20; Length 686;

Best Local Similarity 99.8%; Pred. No. 3e-267;

Matches 653; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 69 CTACCATGTCCTCTTGGAGCAGACAGCGACCAAAAAGCCCGGGGGCATTCAACCCCATG 128  
|||||  
Db 22 CTACCATGTCCTCTTGGAGCAGACAGCGACCAAAAAGCCCGGGGGCATTCAACCCCATG 81  
  
Qy 129 TTTCTAGAACTCTGTTCTGCTGCTGCTGTTGGCAGCCTCAGCCTGGGGGGTCACCTGA 188  
|||||  
Db 82 TTTCTAGAACTCTGTTCTGCTGCTGCTGTTGGCAGCCTCAGCCTGGGGGGTCACCTGA 141  
  
Qy 189 GCCCCAAAGACTGCCAGGTGTTCCGCTCAGACCATGGCAGTCCATCTCCTGTCAACCAC 248  
|||||  
Db 142 GCCCCAAAGACTGCCAGGTGTTCCGCTCAGACCATGGCAGTCCATCTCCTGTCAACCAC 201

Qy 249 CTGCCGAAATCCCGGCTACCTGCCAGCCGACACCGTGACCTGGCCGTGGAATTCTTCA 308  
 |||  
 Db 202 CTGCCGAAATCCCGGCTACCTGCCAGCCGACACCGTGACCTGGCCGTGGAATTCTTCA 261

Qy 309 ACCTGACCCACCTGCCAGCCAACCTCCTCCAGGGCGCCTCTAAGCTCCAAGAATTGCACC 368  
 |||  
 Db 262 ACCTGACCCACCTGCCAGCCAACCTCCTCCAGGGCGCCTCTAAGCTCCAAGAATTGCACC 321

Qy 369 TCTCCAGCAATGGGCTGGAAAGCCTCTCGCCCGAATTCTGCGGCCAGTGCCCGAGCTGA 428  
 |||  
 Db 322 TCTCCAGCAATGGGCTGGAAAGCCTCTCGCCCGAATTCTGCGGCCAGTGCCCGAGCTGA 381

Qy 429 GGGTGCTGGATCTAACC CGAAACGCCCTGACCGGGCTGCCCCGGGCCTCTTCCAGGCCT 488  
 |||  
 Db 382 GGGTGCTGGATCTAACC CGAAACGCCCTGACCGGGCTGCCCCGGGCCTCTTCCAGGCCT 441

Qy 489 CAGCCACCCTGGACACCCTGGTATTGAAAGAAAACCAGCTGGAGGTCCTGGAGGTCTCGT 548  
 |||  
 Db 442 CAGCCACCCTGGACACCCTGGTATTGAAAGAAAACCAGCTGGAGGTCCTGGAGGTCTCGT 501

Qy 549 GGCTACACGGCCTGAAAGCTCTGGGGCATCTGGACCTGTCTGGGAACCGCCTCCGGAAC 608  
 |||  
 Db 502 GGCTACACGGCCTGAAAGCTCTGGGGCATCTGGACCTGTCTGGGAACCGCCTCCGGAAC 561

Qy 609 TGCCCCCGGGCTGCTGGCCAACCTCACCTCCTGCGCACCTTGACCTTGGGGAGAACC 668  
 |||  
 Db 562 TGCCCCCGGGCTGCTGGCCAACCTCACCTCCTGCGCACCTTGACCTTGGGGAGAACC 621

Qy 669 AGTTGGAGACCTTGCCACCTGACCTCCTGAGGGGTCCGCTGCAATTAGAACGGC 722  
 |||  
 Db 622 AGTTGGAGACCTTGCCACCTGACCTCCTGAGGGGTCCGCTGCAATTAGAACGGC 675